IMPROVING THE ANALYST AND DECISION-MAKER'S PERSPECTIVE THROUGH UNCERTAINTY VISUALIZATION

THESIS

Evan T. Watkins, Captain, USAF

AFIT/GCS/ENG/00M-24

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCS/ENG/00M-24

IMPROVING THE ANALYST AND DECISION-MAKER'S PERSPECTIVE THROUGH UNCERTAINTY VISUALIZATION

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Computer Science and Software Engineering

Evan T. Watkins, B.S. Computer Science

Captain, USAF

March 2000

AFIT/GCS/ENG/00M-24

IMPROVING THE ANALYST AND DECISION-MAKER'S PERSPECTIVE THROUGH UNCERTAINTY VISUALIZATION

Evan T. Watkins, B.S. Computer Science
Captain, USAF

Approved:

2 MAR 2000	
Dr. Gregg H. Gunsch Committee Chair	Date
2 Mar 2000	
Timothy M. Jacobs, LtCol, USAF Committee Member	Date
2 MAR 2000	
Dr Robert G. Eggleston Committee Member	Date

Acknowledgments

In the course of this research, many people contributed their time, energy, and expertise to ensure its success. I would like to thank all those who provided support, suggestions and help in one way or another.

I had the pleasure of meeting and working with several folks one of whom was Dr Eggleston of the AF Research Laboratories, Human Performance Directorate. Dr Eggleston you are a great and insightful person; your support makes research like this interesting and challenging.

Dr Gunsch you are enlightening and it was a pleasure to have you as my thesis advisor. Your time, dedication, and critical thinking have made this project a success.

I would also like to thank my sponsor, the National Air Intelligence Center, in particular Captain Bobo and Mr Homer for their assistance with this thesis.

I also want to provide special thanks those close to me for their help. Julie, your help is greatly appreciated; you lifted a load off our shoulders - Thanks. I cannot forget our pets Rusty, Buffy, Cody, and Sammy who were always there, ready to entertain and extend their unbiased gratitude and happiness. As always, a special thanks goes to my wife, Stephany. Stephany, I thank you for your patience, tolerance, assistance, and understanding.

This work would not have been possible without all of you – Thank you.

Evan T. Watkins

Table of Contents

List of Figures

List of Tables

AFIT/GCS/ENG/00M-24

Abstract

Human decisions and predictions are based upon facts, patterns, and intuition, each of which is affected by uncertainty. Facts and patterns change while intuition, unique and disparate among people, is inherently difficult to duplicate or model. However, it is possible to facilitate and enhance decision-making and intuition by providing the decision-maker with better information. Decision-makers and their support require tools that improve the speed and quality of information comprehension, which improves their decisions and reduces error and loss. This thesis explores the practicality of enhancing the information used in decision support systems by including uncertainty without additional information overload.

In this research, I establish the Taxonomy of Uncertainty from the numerous reasons and causes for uncertainty. The taxonomy is used to foster an approach to visualizing the uncertainty associated with an object and existing throughout the decision support and intelligence gathering systems. The resulting approach to including uncertainty involves recording uncertainty, identifying the relevant items, computing and visualizing uncertainty, and finally, providing interaction with the selection of uncertainty.

A prototype that modeled part of the decision support system DIODE was created to embody most aspects of the approach to including uncertainty and was used to validate these efforts. Evaluation responses from several analysts support the thesis that the analyst and decision-maker's knowledge is enhanced with superior and enlightening information afforded by including and visualizing uncertainty, which can improve the decision-making process.

Glossary

Accuracy – the degree to which the result of a calculation or measurement approximates the true value

Acquisition – the process of gaining possession of something; acquiring, learning, or gathering process

Age – length of time something has existed; duration of existence

Ambiguity – when something may be interpreted in more than one way

Bias – a particular tendency or inclination; prejudice; influence, often unfairly; an oblique or diagonal line [Web97]

Completeness – having all normal and necessary parts

Contradiction – is the state of being in opposition; opposing; negation

Credibility – the quality, capability, or power to elicit belief

Deterministic – when repeated trials give exactly the same result

Distortion – a misrepresentation of the actual event, information, or object

Error – a deviation from what is correct, right, or true; mistake

Exposition – the act of exposing, to set forth meaning or intent

Gnomon – is any kind of pointer that indicates a value by casting a shadow, "NO-mun" [Lea99]

Hedges – as applied to fuzzy logic, hedges are terms that modify other fuzzy sets (e.g. very, somewhat, and slightly)

Heuristic – rule of thumb; problem solving technique that leads to a correct solution, but not necessarily the best solution nor the best performance

Historically – based upon previous events

Ignorance – the condition of being uneducated, unaware, or uninformed

Incompleteness – when or where some thing is missing

Inconsistency – lacking consistency, predictability; when there is more than one plausible solution

Incorrectness – when or where the information is wrong

Insoluble – means the unfathomable, undecipherable, unsolvable, and describes situations or information that are unknowable

Intelligence – is the ability to think rationally, act purposefully, and deal effectively with the environment

Language – is a system of combining arbitrary symbols to produce meaningful statements, we interpret words through the use of perception

Limitation – a shortcoming or defect; restriction; constraint; hindrances

Misuse – improper or incorrect use

Non-deterministic – when repeated trials do not give exactly the same result, random

Precision – the exactness with which a number is specified; the number of significant digits with which a number or measurement is expressed; the finer details

Qualitative – pertaining to quality, related to fuzzy logic it is the use of terms to express quality and value (e.g. tall, quick, accurate)

Quantitative – pertaining to a quantity or number, includes numbers, intervals, and linguistic quantifiers (e.g. 5, [0, 0.23], some, few, none, all)

Random – when a value/occurrence fluctuates about the mean without settling to a specific value

Salient – prominent or conspicuous; leaping or jumping; projecting or pointing outward [Web97]

Symbolism – the practice of representing things with symbols or of attributing symbolic meanings or significance to objects, events, or relationships [AHD98]

Symbology – the study, use, or interpretation of symbols or symbolism [AHD98]

Task complexity – refers to difficulty or simplicity for the subject to understand the task

Thinking – involves manipulating mental representations of information to draw inferences and conclusions, often with a goal or purpose

Undecidability – comes from undecided and means indecisive, irresolute, vacillating; possibly because the problem is thought to be insoluble (but is NOT) or because the validity or verifiability is not pertinent (fantasy & fiction) [Smi89]

Unreliability – inability of being relied on; undependable

Vague – not clearly expressed, inexplicit, indistinct

IMPROVING THE ANALYST AND DECISION-MAKER'S PERSPECTIVE THROUGH UNCERTAINTY VISUALIZATION

1. Introduction

> *War is the province of uncertainty: three-fourths of those things upon which action in war must be calculated, are hidden more or less in the clouds of great uncertainty.*
>
> Carl von Clausewitz
> *On War*

1.1 Background

Computerization and automation augment and even replace many aspects of daily human activity. The rise of the Information Age and the global dispersion of computers and sensors provide the tremendous availability and flux of information – information overload is common. In fact, information overload affects every facet of a typical person living in the Information Age; decision-makers, information analysts and information operators are not excluded from information overload. Decision-makers routinely plow through torrents of related and unrelated information to aid their evaluation and response to a situation, also known as the OODA loop [Fad95]; observe, orient, decide and act (OODA).

Human decisions and predictions are based upon facts, patterns, and intuition, each of which is affected by uncertainty. Facts and patterns change while intuition, unique and disparate among people, is inherently difficult to duplicate or model. However, it is possible to facilitate and enhance decision-making and intuition by providing the decision-maker with better information. Decision-makers and their support require tools that improve the speed and quality of information comprehension, which improves their decisions and reduces error and loss. This

thesis explores the practicality of enhancing the information used in decision support systems (DSS) by including uncertainty without additional information overload.

Scientific and medical industries have demonstrated the enormous benefit of information visualization [MDB87]. Through appropriate rendering, substantial amounts of data can be visualized to take advantage of human visual pattern detection and relationships [Kel93, SML97, Tuf97]. Information visualization is already helping decision-makers and operators in various systems, such as the Global Command and Control System (GCCS), Global Combat Support System, and intelligence decision support tools (IDST). IDSTs can be composed of any tool used to aid the intelligence analyst. Information visualization helps by correlating friendly force positions and status against related Geographical Information Systems that include topographical and political maps. The next practical improvement is to visualize uncertainties about those objects, information, and relationships. Uncertainty visualization is a technique of extending the information displayed by including the uncertainty about the data.

There are terabytes of information, centuries of lessons, and hundreds of tools that aid planners and decision-makers, yet very little that depicts uncertainty; the risk of knowing too little. Current visualizations of uncertainty are limited; recent research has extended the library beyond older methods that simply identified numeric intervals. These ranges indicated the range of occurrences or the inability to refine the data to a specific number. Tools like Joint Operations Visualization Environment (JOVE) and the Common Operational Picture (COP) are being developed to aid the decision-maker, yet neither has a representation for the uncertain and both require the knowledge and methods for depicting ambiguous, unclear, unreliable or old information [Ack98, Kor97, DG00]. In the mid 1990's the Chief of Staff of the Air Force commissioned some future-thinkers to identify the technology requirements we might have for concepts and capabilities of 2025. These leading edge representatives had varying visions of 3D, real-time, holographic collaborative tools capable of producing a view of the battlefield and

various planning scenarios [MFJ96, Osb96]. However, the 2025 evaluators failed to consider uncertainty, particularly of the information used to render the visualizations – it is possible that an underlying assumption was that the information used was guaranteed to be factual.

1.2 Problem

Information overload is a common and regrettable side effect of today's information technologies. Users can be inundated by way of technology at a moment's notice or upon request with terabytes of data, much of which is irrelevant, poorly organized, or improperly represented. Most cable TV services consist of almost fifty channels of shows, reports, and commercials running 24-7 – about 99% are irrelevant since the average adult watches about two hours of TV a day [Har97]. Inundation even occurs when searching for specific information on specialized web sites; a June 1999 search for "computer visualization" on the Association of Computing Machinery electronic library web site reveals 18 titles and 479 textual hits. Information is knowledge, yet too much is burdening and useless if not presented in a manner that makes translation and inception practical and speedy. Can the quality of information in current DSS be increased by augmenting the systems to include indicators of uncertainty thereby improving a leader's OODA loop?

The overarching objective of this thesis is to provide a decision-maker with decision-making assistance by including and identifying uncertainty existing throughout the decision support system. Information overload, holes and conflicts in the data, side effects of actions, implications of decisions and so forth impede and flaw perceptions as well as decisions. How can we let the decision-makers know there is risk or uncertainty in a particular object or information that may affect their decision without information overload?

To address these problems, the focus of this research is to establish a taxonomy of uncertainty and methods of visualizing uncertainties thereby providing the decision-maker with

more and better information on which to base his decisions. The investigation will provide information and recommendations that are expected to improve the decision making process, essentially the OODA loop. Research will also address issues related to the human factor side of information visualization: what should be visualized and how to depict or quantify uncertainty such that it helps the user grasp the nature or existence of the uncertainty.

1.3 Scope

The specific target of this research is to improve any IDST and DSS. A portion of the Dynamic Information Operators Decision Environment (DIODE) will be reproduced in a prototype to demonstrate the viability of the concepts presented in this work. The improvements will be accomplished by categorizing numerous reasons and sources for uncertainty to create an uncertainty taxonomy. The taxonomy will be used to develop an approach to visualizing the uncertainty associated with objects used in and existing throughout the DSS and intelligence gathering systems. The taxonomy provides an understanding of the categorization of uncertainty and will bolster the visualization of the uncertainties. The methods for depicting uncertainty are applicable to system designers considering ideas and approaches for including and visualizing uncertainty. Although most of the research focuses on any IDST, the methods and taxonomy are not restricted to any one particular system and intentional abstraction facilitates supplementing any visual tool.

1.4 Assumptions

The following assumptions are precursory and necessary to perform the research particularly due to the fact that some issues are beyond the scope of this thesis. Several declarations are included among the assumptions. First and foremost, the information required to reveal uncertainty is preprocessed, stored, and available for reference in a database, and the

supporting information is also available – the research conducted will not include data retrieval, inference, or attribute calculations. The process of building or adding to the database is immaterial; software agents and analysts might be responsible for the data. The basic assumption is that various attributes are available and include information about the uncertainty contained in the data that can be scaled, graphed, or rated (high, medium, or low) and may include explicit numbers. Second, the symbols used in any military application comply with those contained and represented in MIL-STD 2525[1]; therefore, representations of uncertainty will be based upon these symbols and should be applicable to unforeseen additions. The representations should enhance or augment the standard symbols, not replace them. The third and final assumption is that the DSS visualizations are dynamically rendered. This makes it possible to add to or augment the current depiction.

1.5 Approach

This thesis and research will be accomplished through several tasks and can be categorized into three rudimentary steps: research, define, and analyze. These steps represent significant segments of this thesis and are not independent of the other processes and considerations. During the research phase, I will investigate cognition, decision-making, uncertainty, and information visualization, establishing the foundation for understanding information visualization, uncertainty, and reasoning with uncertainty. This foundation is used to define a taxonomy of uncertainty and the categorization of uncertainties. By identifying the taxonomy, I will be able to define a strategy for visualizing uncertainty in an IDST. Finally, applying the proposed methods to a prototype and analyzing the results validates the research.

Various visualization techniques will be examined, relevant human computer interface (HCI) issues will be identified, and cognitive issues considered. The current state of visualizing

[1] MIL-STD 2525 is the DOD guide to NATO Warfighting Symbology.

uncertainty and DSS will be investigated. In addition to visualization and cognition, uncertainty and reasoning with uncertainty will be explored to increase general understanding and insight to potential relationships.

The uncertainty taxonomy will be assembled from other related expert materials if one does not already exist and will include an exposition of its composition. There will be an itemization of uncertainty visualization techniques that might be applied to the approach I define for presenting uncertainty in an IDST. The thesis will include explanations of the options, issues to avoid, considerations, as well as assumptions. The proposed methodology will be provided to the sponsor, the National Air Intelligence Center, for evaluation.

Analysis will include a prototype program that demonstrates uncertainty visualization in the IDST and the improvement over current tools and systems. A demonstration of the program followed by evaluations, surveys and responses to the visualization of uncertainty will be used to indicate successful augmentation.

1.6 Thesis Organization

Five chapters construct the thesis. Chapter 2 examines a great deal of background material covering three central topics: cognitive issues, uncertainty, and information visualization. Sections 2.1 through 2.1.3 cover the cognitive concerns examined as considerations in design. Those sections acknowledge and summarize several cognitive and decision-making strategies. The heuristics and biases that affect decision-making are presented in the Section 2.1.2. The Taxonomy of Uncertainty is developed in Section 2.2, which provides the basis for the hierarchy as well as the contributing sources. More specifically, Section 2.2.4.2 examines ignorance and Section 2.2.5 presents the completed Taxonomy of Uncertainty. Chapter 2 also covers information visualization, techniques and goals. A brief synopsis of several common uncertainty visualization techniques is provided in Section 2.3.4; Section 2.3.4 and its

subsections cover the uncertainty visualization issues, goals, and options that made up some of challenges of this thesis.

Chapter 3 explains considerations addressed while conceiving this approach to visualizing uncertainty in a DSS. It covers four central themes: appreciation for the generic term "object," including uncertainty in DSS, some ideas for estimating multi-dimensional uncertainty, and a description of the prototype program. The generic "object" and its diverse classification of information is discussed in Section 3.1. Section 3.2 elaborates on the approach for including uncertainty in a DSS, which is founded on identifying and presenting uncertainty. Section 3.3 presents a few ideas for estimating multi-dimensional uncertainty and what it means to be "high in uncertainty." The model program and demonstration are described in Section 3.6.

The prototype program was evaluated by several of the sponsor's intelligence analysts; Chapter 4 presents the evaluation, results, and analysis of the critiques. Chapter 5 wraps up this work and presents the findings, and recommendations as well as considerations for future projects.

2. *Background and Systems Review*

So far as laws of mathematics refer to reality, they are not certain. And so far as they are certain, they don't refer to the reality.

Albert Einstein

Through a cursory look at some decision support systems (DSS) and tools used by the military I found that none of the environments explicitly included or expressed uncertainty, let alone, even acknowledged that uncertainty was present in the decision-making situation. It was obvious to me that few or no situation ever includes complete knowledge and in other words, is exempt from uncertainty.

Analysts and decision-makers are provided with many sources of information from which they are expected to work and make reasonable decisions from. Each source of information includes some degree of uncertainty and may even introduce uncertainty via the process and medium used to transport the information. Uncertainty can also be introduced at the source as well as in the gathering stages. We employ various techniques to reduce and eliminate uncertainty particularly by increasing the certainty about the information that we do know [Cle96], which indirectly reduces the uncertainty, but never eliminates it. In fact, we can be certain that uncertainty is rarely, if ever, completely eliminated.

It seemed like the current approach for handling uncertainty was to reduce it as much as possible, ignore it, and omit it. I felt the DSS and its users could employ a different approach to handling uncertainty by actually identifying and expressing it since it was an inherent part of the system. By identifying and expressing uncertainty, I also thought it would be possible to include more information with uncertainty.

The goal of this thesis was to find an approach to enhance the information the analyst and decision-maker used in order to improve their perspectives. I expected that the analyst and

decision-maker's perspective could be improved by including uncertainty and uncertain information thereby providing them information that might have otherwise been omitted. This could be accomplished by identifying and visualizing uncertainty and would be facilitated by some technique for visualizing and identifying uncertainty. Through my research, I found support for my ideas in two of four approaches to managing uncertainty featured by Gulick and Martin [GM88]. Gulick-Martin report that we should: (1) recognize and give due attention to uncertainty because making decisions under uncertainty is a fact of life; and (2) communicate the extent of the uncertainty avoiding the suppression of uncertainty.

In order to accomplish my goal I determined out that I needed to explore and understand information visualization, human computer interaction, cognition, reasoning, decision-making, and biases among other things. The plethora of background material included the decision-making processes, the biases and influences affecting decision-making, and additional cognition related information. In addition, I explored issues related to information visualization, learned about various human factors that are considered in human computer interfaces and to avoid persuasive techniques.

I partitioned this chapter into five parts, three of which establish the foundation for this my work. The first section provides a summary of cognition and decision-making related information that I used to better understand the processes and issues involved in decision-making. The second section provides a causal analysis of uncertainty and results in the compilation of a more inclusive Taxonomy of Uncertainty. It also includes a summary of method I chose for identifying and handing uncertainty: fuzzy logic. The third area discusses information visualization as well as uncertainty visualization techniques and includes a small section on symbology. This section also provides some insight to the systems I looked at to determine their use of uncertainty. The fourth section introduces the programming package I used to create a prototype program. Finally, the chapter is wrapped up in the fifth section.

2.1 Cognitive Issues

Humans are imperfect; the following sections provide a brief overview of information I felt was critical to recapping some aspects of decision-making and the different issues that affect decisions.

2.1.1 Cognition

Cognition refers to the mental activities involved in acquiring, retaining, and using knowledge. The fundamental cognitive processes are perception, learning, and memory; thinking, language, and intelligence are manipulations of mental representations of information rather than fundamental processes [Hoc97].

We are human and have faults that affect our cognition as well as decision-making. One of our greatest limitations is our inability to handle large amounts of information in our limited short-term memory. Miller found that people are constrained in the number of items they can keep active in memory: we are limited to seven plus or minus two items [Mil56]. This cognitive limitation greatly influences our performance, the variety of tasks we can accomplish, and decision-making.

Our visual senses and cognitive manipulations provide for perception, identifying what we see or think we see, from which we infer and interpret additional information such as knowledge and memories. Our cognitive limitations, particularly our inability to handle large amounts of information, lead us to incorporate various mechanisms and strategies into our being so it is possible to deal with complex situations, events, and decision-making situations. These strategies are facilitated by a number of heuristics that can be tainted by a variety of biases. The following sections present some cognitive strategies, heuristics, and biases.

2.1.1.1 Coping Mechanisms and Decision-Making Strategies

Humans incorporate several coping mechanisms to deal with our cognitive limitations. Dahl examined cognitive issues relative to military command and decision making, identified four coping mechanisms: *editing* or *eliminating by aspects*, *decomposition*, *pattern recognition*, and *framing* [Dah96]. Hockenbury, a psychology researcher, presents these and two other ideas, *singling out* and *accumulating*, as decision-making strategies used to cope with complicated situations of dealing with difficult decisions [Hoc97].

- *Eliminating by aspects* (and editing) is when a person filters problem data before reaching an essential point. This occurs when people consider and eliminate potentially unnecessary information and alternatives before analyzing additional information.

- *Decomposition* occurs when we break a large problem into component parts. This scheme can be ineffective if the sub-problems are independently insoluble. This can also increase your cognitive load rather than identifying simpler components.

- *Pattern recognition* is our search for recognizable patterns; we can and tend to use previous experience when dealing with familiar situations.

- *Framing* is the human ability to put a situation or problem into a general set of beliefs and perspectives that constrain data collection and analysis thereby narrowing the information search and association.

- *Singling out* occurs when we base our decision on a single feature. This typically occurs with simple decisions but as humans, we are not free from using this strategy in any situation.

- *Accumulating* factors or features is a strategy that increases in value based upon the amassing of perceived or intended values of each alternative. The strategy builds evidence supporting a particular decision.

These six points identify the different ways a decision-maker might examine a situation and come to a conclusion. Decision analysis (DA) is a more formal approach for helping people with difficult decisions. DA provides some techniques, structure, and guidance for decision-making, but does not tell people how to make decisions. Several tools to aid decision-making have been developed [Cle96]; some include implementations of Bayes theorem, Bayesian network, fuzzy logic, as well as other models that are covered in Section 2.2.6.

2.1.1.2 Cognitive Rules and Strategies

Dahl also discusses our (human) ability to generalize some requirements in problem solving situations [Dah96]. We use cognitive rules and strategies like *satisficing*, *analogizing*, *incrementalism*, *blurring with probabilities*, and *nutshell briefing* to work through perplexing situations.

- *Satisficing* is when we accept a solution that is 'good enough' rather than solving for the optimal result, which may be unreachable.
- *Analogizing* occurs when we seek similarities to a situation through our comparison of other situations we have experienced or learned, occasionally and erroneously ignoring vast differences.
- *Incrementalism* is the practice of making small changes to lighten the load of the overall situation.
- *Blurring with probabilities* is the tendency to misuse or misinterpret statistical data to explain complex events or sustain preconceptions.
- *Nutshell briefing* is the summary of a situation or event into a succinct "in a nutshell view" usually provided by support staff upon which the decision-maker uses for his decision.

These five points identify strategies that can be shortcuts as well as problems. The *blurring with probabilities* strategy seems to be the least accurate and least appropriate for decision-maker use. I also think that the *nutshell briefing* strategy can be inappropriate when the decision-maker uses only the information provided by that briefing to make a decision. Although pervasive in large organizations, e.g. the Department of Defense (DOD), I would hope that a decision-maker would not act without first having trust in his support staff or having other information that supports their input.

2.1.1.3 Decision-Making Styles

Sauter provides a different perspective on reasoning and decision-making in her recent paper "Intuitive Decision-Making." Sauter identifies four types of decision-making styles: *left-brain*, *right-brain*, *accommodating*, and *integrated* [Sau99].

- The *left-brain* style employs rational reasoning stressing analytical and quantitative techniques, sequentially applying logic and data to resolve sub-problems decomposed from a greater one. This style works best with complete information and when "relevant variables can be controlled or predicted, measured, and quantified." Frequently, the conditions and information for this style of decision-making are not present, thereby preempting analytical methods. In addition, immeasurable factors such as values and morals are not addressed.

- The *right-brain* approach places more value on feelings than facts using intuitive techniques. The brainstorming and emergent trend projection are appropriate uses of this style. There is an "unstructured and spontaneous procedure of considering the whole rather than its parts," even with insufficient information. The problem with this method is the lack of reproducibility and provable theories. See Section 2.1.1.4 for a breakdown of intuition.

- The *accommodating* style is used when experienced decision-makers realize certain situations call for a style opposed to their dominant strategy. They employ an alternative style during these situations.

- The *integrated* approach attempts to take advantage of the benefits involved with both left and right-brained styles. Using intuition with analytical processes allows the decision-maker to address immeasurable issues as well as uncertain and complex elements. See Section 2.1.1.4 for a breakdown of intuition.

Sauter's work is different from Hockenbury's strategies and Dahl's coping mechanisms by her association of styles to the left (analytical) or right (creative) sides of the brain as well as her recognition and inclusion of intuition.

2.1.1.4 Intuition

Intuition is difficult to define and explain, yet exists and is required in many decision-making environments whether or not it is recognized. Sauter identifies intuition and intuitive thought as an escape from being constrained by her categorization of decision-making styles. She explains that in addition to avoiding a particular strategy, intuition can be used to address many uncertain, complex and immeasurable elements in a practical manner [Sau99].

American Heritage Dictionary defines intuition as the act or faculty of knowing or sensing without the use of rational processes. Sauter uses several sources to elaborate on and define intuition. She essentially explains that intuition refers to a "sense of feeling of pattern or

relationships, immediate insight, a sudden awareness, knowing an answer without knowing how it was reached," integrating dissimilar chunks of knowledge, and sensing "patterns among unrelated facts."

Forms of intuition, classified by the perceived method of triggering, are listed below.

- *Illumination* is the sudden awareness
- *Detection* is the revelation of facts or answers when working on another problem
- *Evaluation* is the feeling of confidence or "what feels right" when provided with choices
- *Prediction* involves hypothesizing without evaluating the data
- *Operative intuition* provides a sense of direction, suggesting something requires another look or exploration
- *Creative intuition* involves options and possibilities often supplementing detection by generating other ideas

Sauter wraps up intuition by explaining that some people become intolerant of details and routines using intuition as a short cut, ignoring facts and formulas. Intuition can be harmful as well as inventive. The person using intuition in decision-making should therefore be aware of the strengths and weaknesses as well as the potential to erroneously follow inspiration. Lack of experience is a problem affecting some environments, hence others often learn from the experiences of others and their intuition.

This links back to DSS in the sense that some systems can be designed to foster intuition rather than simply report results, but it has to be done carefully so as to prevent abbreviating other potentially formal processes. The DSS should help the decision-makers understand what they know, help them understand the underlying assumptions, and help them test assumptions or intuition.

2.1.2 Heuristics and Biases

Kahneman and Tversky made enormous contributions to reasoning and cognitive sciences with their various studies, papers, and evaluations. One such writing discusses our decision-making processes and the heuristics used to handle the uncertainty [KT82] that may be associated with various decisions. The description of several heuristics people use to assess probabilities and predict values to aid decision-making are provided below with the various biases can affect these heuristics. Table 1 summarizes the heuristics and biases presented.

The heuristics employed to assess probabilities and predict values are *representativeness*, *availability*, and *adjustment and anchoring*. These heuristics identify the methods people may employ in decision-making situations in order to accomplish previously mentioned decision-making strategies. The biases affecting these methods are also provided as well as some ways to avoid the misguided influence. By understanding the heuristics, we can understand what to expect from a decision-maker and how to reduce improper biases in order to increase the likelihood of making a better decision.

Table 1. Kahneman-Tversky Heuristics and Biases

Heuristic	Biases
Representativeness	Insensitivity to prior probability of outcomes
	Insensitivity to sample size
	Illusion of validity
	Misconception of chance
	Insensitivity to predictability
	Misconception of regression
Availability	Retrievability of instances
	Imaginability
	Illusory correlation
	Effectiveness of a search set
Adjustment and anchoring	Insufficient adjustment
	Evaluation of conjunctive and disjunctive events

2.1.2.1 Representativeness

Many of the probabilistic questions we deal with attempt to identify what we think a probability might be, its likelihood of originating from another process, and the probability of causing another event. Sample questions of this nature include: What is the probability that object A belongs to class B?, What is the probability that event A originates from process B?; and What is the probability that process B will generate A?

Kahneman-Tversky found that when people answer these questions they typically rely on the "representativeness" heuristic. This heuristic occurs when probabilities are evaluated by the degree to which A is "representative" of B; in other words, the degree to which A resembles B. The probability that A originates from B is assessed high when A is highly representative of B. On the other hand, if A is not similar to B, the probability that A originates from B is judged to be low.

The biases associated with representativeness are the *insensitivity to the prior probability of outcomes*, the *insensitivity to the sample size*, the *misconception of chance*, the *insensitivity to predictability*, the *illusion of validity*, and the *misconception of regression.*

- **Insensitivity to prior probability of outcomes.** Kahneman-Tversky hypothesized and found that if people calculated probabilities by representativeness they would ignore prior probabilities and base-rate frequencies. However, they also found that prior probabilities are used correctly when no specific evidence is provided.
- **Insensitivity to sample size.** People tend to ignore the sample size and assume the likelihood of obtaining a specific result for the entire population is inherited by smaller samples of that population. Consequently, when using representativeness, the determined probability of a sample statistic is incorrectly independent of the sample size.
- **Misconception of chance** [and randomness]. People erroneously expect random sequences from random processes also known as "the gambler's fallacy." For instance, people tend to believe that flipping a coin will result in very random sequences. They believe the sequence HTHTTH is more likely to occur than HHHTTT because it is more random and more representative of the expected sequence generated by a random process.

- **Insensitivity to predictability**. Errors also occur when people use representativeness to make numerical predictions. People tend to use the description of team or object to judge how well it will perform. Predictions based upon the favorableness of the description versus information relevant to the performance are incomplete and inaccurate.

- **Illusion of validity**. The representativeness heuristic we use tends to cause us to select an outcome that is most like the input and our stereotypical perceptions. However, we are less likely to change that prediction when there is the impression of a high degree of likeness between the outcome and input, hence the illusion of validity.

- **Misconception of regression**. A concept overlooked and misunderstood by many people is the phenomenon of regression towards the mean in consecutive examinations. Regression toward the mean is described as the natural tendency for comparisons, measurements, or examinations taken consecutively to change or move closer to the average of a global sample.

2.1.2.2 Availability

When we are asked to predict an event or assess a probability, we typically determine that answer by similar instances and occurrences that can be recalled and the ease that they can be retrieved. For instance, my perception of heart disease or cancer limited because I know few people suffering with these problems.

On the other hand, availability can be useful when determining frequency and probability, because "instances of large classes are usually reached better and faster than instances of less frequent classes" [KT82]. Factors other than frequency and probability lead to several biases affecting availability predictions. The biases associated with availability are the *retrievability of instances*, the *effectiveness of a search set, imaginability*, and *illusory correlation.*

- **Retrievability of instances**. People commonly estimate the likelihood of an event or the frequency of occurrence based upon information they recall. The more familiar or salient an event or information is, whether through details or emotions, the more likely people believe it will occur again. This retrievability causes errors in estimation and judgement.

- **Effectiveness of a search set**. We use our memories and whatever we can recall quickly to create a sample space by which to make some not-so-simple decisions. Our ability to identify many items from memory that fit one answer does not mean

that our search set is accurate; in fact, our memories are not as effective as the true solution space.

- **Imaginability**. Sometimes we need to assess a probability or frequency of something we do not know about, but we have some information or rules that can shape our imagination. People tend to imagine situations they believe are representative of actual occurrences, then they estimate the probability by how easy it was to envision the events or create the scenario.

- **Illusory correlation**. When people assess joint or conditional probabilities that depend on the correlation or connection between two events, they may misinterpret the co-occurrence of two events as the strength of their association. If the associative bond between two events is very strong or perceived as such, then it is easy to conclude that the events occur more frequently than they do in reality. This illusionary correlation can also interfere with the detection of other relationships or events that are present.

2.1.2.3 Adjustment and Anchoring

People often select a prominent starting point and adjust their guesses or theories when they make predictions. Anchoring is the phenomenon in which different starting points result in different assessments that are biased towards the initial starting values. Biases related to anchoring and adjustment include *insufficient adjustment* and the *evaluation of conjunctive and disjunctive events*.

- **Insufficient adjustment**. Kahneman-Tversky's research shows that people tend to make insufficient adjustments to starting points.

- **Evaluation of conjunctive and disjunctive events**. Whether by misunderstanding of probability or the misunderstanding of the problem, people tend to be biased in the evaluation of disjunctive and conjunctive probabilities, underestimating the former and overestimating the latter. Apparently, the chain-like descriptions of conjunctions lead to overestimation, while the funnel-like structure of disjunction lead to underestimation. Overestimating the probability of conjunctive events leads to unwarranted optimism.

2.1.2.4 Biases Identified by Alternative Sources

Several other influences and biases identified in other studies are discussed below, in no particular order.

Hindsight bias. "Hindsight is 20/20" and some people can not believe in any other outcome once they have rationalized the "only possible outcome." Fischhoff's research revealed that when people become aware of an outcome they reinterpret earlier knowledge and perspectives making sense out of the "more likely result" [Fis82a]. When this happens people also tend to discount other possibilities and remain focused on the expected inevitable result. Once results are known people tend to assume they were the only outcomes that could have happened and underestimate other potential outcomes. Fischhoff has also found that foreknowledge and warnings of hindsight bias have little effect.

Conservatism. People are generally conservative in their calculations of information [Edw82]. One of the main reasons for this conservatism is our basic misunderstanding of the problem, diagnostic meaning and the ensuing miscalculations. Consider the simple example of two bags of marbles: one with 70 red and 30 blue marbles, and the other with 30 red and 70 blue marbles. Suppose that a bag is chosen at random and a marble drawn. Suppose the marble is red, then it is replaced. After twelve draw-and-returns, you count 8 red and 4 blue marbles. What is the probability that this is the bag with 70 red marbles? Most people would have answered between 0.7 and 0.8; probably guessing based on the number of red marbles that occurred. The proper formulation uses a form of Bayes' theorem; the answer is 0.97.

Overconfidence. Oskamp repeated research that tested the confidence and accuracy of some psychologists diagnosis mental illness [Osk82]. The test subjects' confidence soared as they received more information, but their accuracy changed very little; the certainty about their own decisions became completely out of proportion to the correctness of their decisions. The study showed that confidence increases as more information is acquired and that most people are overconfident about their judgements. Clearly, the tendency to be overconfident is worth noting; as we gather more information about a situation or decision, our increasing confidence is not an accurate indicator of our increasing accuracy.

Overestimation effect. This is our tendency to overestimate the rarity of some events, another form of misunderstanding probability and chance [Osk82]. For instance, it would be common to determine that it is virtually impossible for two people in a class of 23 students to have the same birthday. However, the odds are about one out of two, 50-50 [Osk82].

Belief-bias effect. Matlin identified that people sometimes accept only the evidence that conforms to their belief and reject or ignore anything to the contrary [Mat89]. People often stubbornly refuse to be convinced of anything other than what they have set their mind to be factual. The example cited involved a study where believers of ESP were steadfast in their conviction after witnessing several ESP communication trials designed to convince them that ESP was phony.

Confirmation. This bias is the strong tendency to seek out evidence that affirms a belief, while avoiding information that discounts or disproves the belief [Mat89]. For example, some people will read articles and editorials that support or interpret events with a similar perspective and avoid those that are different or disagreeable.

Fallacy of positive instances. This is our penchant to recall rare events that seem to confirm our beliefs, while ignoring the coincidence [Mat89]. For example, we tend to remember the seemingly special or impressive events such as when if you happened to be thinking of a friend right before they called versus the many times this did not occur. Although not a major behavior influence, it is a bias that impedes proper decision-making.

2.1.2.5 Other Factors

Many other issues and factors affect our rationality, decision making, probability assessment, and analysis. These include task complexity, response modes, list lengths, measurement scales, and even phobias. Physical differences or weaknesses can be considered a

factor. It is estimated that approximately nine percent of the human population suffers from some genetic color vision deficiency [Lev97].

Fortunately, some physical deviations can be accommodated for, while other problems can not be as easily addressed. Focus and attention span is an issue, some people do not concentrate enough on easy tasks; when they perceive the task complexity is low, they essentially get lazy. Response modes are the methods of responding to actions and were shown to influence elicited probabilities; people provide different values when asked to assess fractiles, odds, or quantities. List-length affects the probabilities that people attach to various events. They can even be influenced by the events listed for consideration. With a list of n alternatives some people will anchor on the simple probability of $1/n$ and then adjust insufficiently. Measurement scale refers to the scale in which the evaluations or quantities are expressed. Some people are more comfortable thinking and expressing in terms of one scale over another, such as the binary rather than hexidecimal. Phobias can influence routine activities and will even affect rational decision-makers. Consider the number of buildings that eliminated the perception of the thirteenth floor because of the "bad luck associated" with the number thirteen.

2.1.3 Cognitive Issues Summary

The previous sections provide some insight into our own decisions and the possible effectors of those decisions. Dahl identifies several coping mechanisms that Hockenbury presents as decision-making strategies. Both explain techniques that we employ to help with difficult decisions and our cognitive limitations. I also provided Sauter's different perspective to decision-making: she identifies four decision-making cycles, two of which are compromises between the left and right brain styles. Sauter also explains that intuition can help in situations where the decision-maker does not want to commit to a particular strategy and can be practical

for addressing many uncertain, complex, and immeasurable elements. Intuition can also be a problem when people become intolerant of details and use it as a short cut.

Heuristics and biases were also discussed. Heuristics refer to problem solving techniques and in relation to decision-making, they are our ways of handling different decision-making situations. Heuristics can be affected by several biases that impede our estimations or judgement. Through understanding, we can reduce improper influences and biases increasing the quality of our decisions.

Although the previous sections only provided biases that were essentially harmful to decision-making, I have not presented the biases that are beneficial to us and our processes. Biases can actually foster pattern recognition, lead us to reasonable theories and estimates, and help us eliminate extraneous information. Bias helped to determine the articles that I chose to include in the background material; otherwise, there would be a ridiculous amount of information most of which repeats.

The information provided in the previous sections was taken into consideration during concept and prototype development. As previously mentioned, I expected to provide a method for including uncertainty that would actually foster intuition and minimize cognitive issues. Understanding the decision-making strategies and the potential for negative biases would help me produce an appropriate solution for identifying and presenting uncertainty.

2.2 Uncertainty

Uncertainty is an essential element to this thesis and is explored and discussed in the following sections. These sections also provide several explanations for uncertainty that are used to establish my own hierarchical organization of the sources of and causes for uncertainty. The hierarchical ordering of the sources leading to uncertainty defines my Taxonomy of Uncertainty used to establish a technique for visualizing uncertainty in a DSS.

2.2.1 Increasing Presence of Uncertainty

In general, we are increasing our understanding and recognition that uncertainty is present throughout our daily lives. Whether or not we are aware of it, we routinely handle details relative to uncertainty, e.g. in weather reports, while traveling to work, and making weekend plans. We often receive, expect, and provide details associated with uncertainty when we exchange information especially when it involves decision-making. Consider the weather reports, they typically include numerous details allowing users of those reports to make well-informed decisions. Imagine the difficulty of making a decision that could be affected by the weather if the report was simplified to one detail rather than several. We might get a report that said "cloudy tomorrow" instead of "mostly cloudy with a 70% chance of rain."

This leads me to the conclusion that general information and predictions without a statement of its uncertainty (or certainty) could be incomplete. The source of information assumes that the user does not need the extra details or that he can infer the uncertainty or lack thereof from the information. Otherwise, the information is truly missing details that could help the user better understand the content particularly when using it for a decision.

As I mentioned, people are recognizing that uncertainty is prevalent and we are increasingly using that information in our decisions. Gulick and Martin point out that the use of uncertainty representation to aid decision-making is increasing, but tapered-off in the 1970s for several reasons. High expectations of Artificial Intelligence (AI) lead to the decrease of uncertainty representation. They specifically cite the cost and low state of computer technology, the resistance of users to quantifying their uncertainty, the difficulty involved with the explicit treatment of uncertainty, and the ideal held by AI researchers that AI systems could resolve the uncertainty. AI researchers believed they did not have to formally deal with uncertainty because of the inherent power of AI: "AI systems could mimic intelligence experts, locate answers in huge solution spaces, or use deductive predicate calculus to discover other facts" [GM88].

2.2.2 *Lacking Uncertainty in DSS*

As Gulick-Martin indicated, I could not find a DSS that represented or expressed uncertainty. I found that many systems lack a information primarily because we avoid putting data and details into the system until we eliminate most or all uncertainty about the information. I noted this observation earlier; it led to the idea for this research. Omitting and delaying information with uncertainty does not represent the real world where we regularly deal with and handle imperfect information. On the other, the analysts and decision-makers using these systems are expected to work without this knowledge let alone the fact that it was omitted. Why do we delay information processing and exchange when some doubt exists? Do we expect to achieve perfection? The Roman scholar Pliny notes that "the only certainty is that nothing is certain" [MSB98].

This thesis engages the failure to include all available and relevant information (in DSS) and proposes an approach for allowing users to include less certain information. This uncertain information would need to be identified and expressed to make the system users aware of its presence so they are not blindly using imperfect data. For the most part, I found that analysts providing data to a system tend to avoid submitting incomplete or inaccurate information in order to avoid the associated uncertainties as well as to provide as much certain data as possible. However, if we can somewhat confidently and clearly identify the cause for concern, we can include that data with the reasons for its uncertainty. Additional information as well as the lack of it affects our perception and understanding; we need a way to include information that is commonly omitted.

2.2.3 *What is uncertainty?*

What is uncertainty and where does it come from? Such a remarkably important, yet difficult question has increased in popularity over the last two decades as people attempt to

improve their understanding of the issues they face during decision-making. While previous approaches and practices attempted to increase our certainty about information, we are now facing and dealing with uncertainty head-on. Uncertainty and its use are often difficult to express and sometimes used poorly. The following quotes are some observations demonstrating the difficulty of grasping uncertainty.

- The uncertainty principle of quantum mechanics specified in 1927 by the German physicist Werner Heisenberg, states that "increasing the accuracy of measurement of one observable quantity increases the uncertainty with which other quantities may be known."
- "Uncertainty refers to imprecision in estimates of future consequences which are conditional on present actions" [Sch96].
- Uncertainty is when something is "not known precisely, not clearly determined, doubt, hesitancy, unpredictability" [Web97].
- "Uncertainty can be considered as the lack of adequate information to make a decision" [GR94].
- "Uncertainty has proven to be the most relevant factor in making decisions. It bounds what the decision maker does and does not know about a decision situation" [GM88].

As the previous quotes indicate, a comprehensive and precise definition of uncertainty was not determinable. I concluded that while defining uncertainty did not help with identifying a technique for expressing uncertainty, ascertaining and organizing the causes leading to uncertainty could be beneficial. Like a previous idea, this one was also supported by the Gulick-Martin findings I discovered in my research.

In light of the difficulty uncertainty causes for decision-making Gulick-Martin identify and summarize several constructive ways to manage uncertainty. They remind us that the principal purpose of intelligence gathering is to remove or reduce uncertainty in decision-making. They also highlight that it is unlikely that uncertainty will be completely removed because uncertainty is introduced in the collection, analysis, and distribution process as well. The four Gulick-Martin approaches to managing uncertainty [GM88] are:

- To recognize and give due attention to uncertainty. Making coherent decisions under uncertainty is a fact of life.

- Understand uncertainty from a substantive, intelligent point of view. What are the sources of uncertainty in the data, devices, sensors, biases, interpretations, filtering, and so forth?

- Apply appropriate analytical tools and techniques to clarify and deal with uncertainties.

- Communicate the nature and extent of the uncertainty. Avoid suppression of uncertainty. Avoid overconfidence. Clarify to fit the user's terms, culture, and jargon.

In hindsight, it is interesting to note the similarities of my idea to their approaches for managing uncertainty. I recognized the failure to include and identify uncertainty that was clearly present in a DSS from which this thesis proposes to develop an approach to enhancing the information used by analysts and decision-makers. The similarities continue as the approach develops. One of the points I reiterate throughout this work is the necessity of expressing and identifying uncertainty. The thesis promotes these uncertainty management techniques through its visualization, processes, and detailed view of the data.

2.2.4 Published Breakdowns of Uncertainty

The ensuing sections identify and evaluate some published views and breakdowns of uncertainty that are used to develop my Taxonomy of Uncertainty. These views include uncertainty, ignorance, and error to more clearly identify the sources of uncertainty prior to pursuing an approach to presenting it. These sections provide pertinent background designations and explanations used to create and extend credibility to my Taxonomy of Uncertainty that is discussed in Section 2.2.5.

I established this taxonomy because I was unable to find a well-defined and studied classification of uncertainty or its causes. Instead, I found several general discussions and only a few hierarchical classifications of uncertainty or that included uncertainty. Most were domain

specific like medical diagnosis represented in Section 2.2.4.3. It became obvious that uncertainty was difficult to nail down to general but useful terms and I would have to establish a classification that met my needs. I thought identifying the fundamental composition of uncertainty was crucial to deriving further explanations of uncertainty and the possible categorizations of sources or causes of uncertainty. I realized later that I could not identify a composition of uncertainty, but I could hierarchically identify terminology for the causes leading to uncertainty. As previously mentioned, I expected to use this hierarchy to aid the identification of an appropriate technique for visualizing uncertainty.

Sections 2.2.4.1 through 2.2.4.3 provide some of the more relevant and useful hierarchies of or including uncertainty that I use to create my Taxonomy of Uncertainty. First, I present the Kahneman-Tversky Variants of Uncertainty to establish the basis for two fundamental classifications of the causes for uncertainty. The second taxonomy is Smithson's Taxonomy of Ignorance, which I use to extend the internal cause for uncertainty that Kahneman-Tversky suggest one reason for it: ignorance. The third taxonomy provided is a break down of the types of error that I use to create an improved categorization of error, which is used as a reason for ignorance. The fourth and fifth sections help identify a term for the other main cause that Kahneman-Tversky identify that uncertainty is attributed to, external factors that I tie to unreliability. Although I start out correlating Kahneman-Tversky's Variants of Uncertainty to my Taxonomy of Uncertainty, it quickly changes as I extend each of the classifications by including many different causes for each. I identify and extend the different portions of the taxonomy as each resource is discussed. They are pulled together as the Taxonomy of Uncertainty in Section 2.2.5 without repeating explanations that were provided as the portions were identified.

2.2.4.1 Kahneman-Tversky Variants of Uncertainty

Kahneman and Tversky define the Variants of Uncertainty based on and involving psychological and phenomenological arguments [KT81]. They establish that uncertainty can be attributed to our *internal* state of knowledge and the *external* world as shown in Figure 1. The outcome of a game and the draw of a card are samples of uncertainties attributed to the external world. Statements like "I think it will rain" and "I hope I remember where I parked" are representative of internal uncertainty. Internal and external uncertainty can also be viewed as events that, respectively, we can or cannot control [KT81].

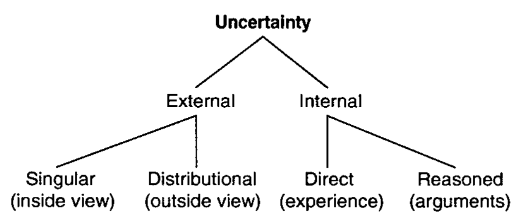


Figure 1. Variants of Uncertainty [KT81]

Kahneman-Tversky use the second layer to identify four modes of judgement people adopt to assess the uncertainty. Internal uncertainty could be determined in a *direct* manner by way of recalling experiences. Internal uncertainty could also be determined through a *reasoned* approach using various arguments for or against something in order to induce an answer from other knowledge [KT81]. External uncertainty can be estimated using *singular* and *distributional* modes of reasoning. The *singular* approach refers to the view gained from one sample or self analysis in terms of what "I might do." The *distributional* mode considers the relative frequencies of an event to determine the uncertainty. In other words, the *singular* mode is what we do when we determine uncertainty based on how we might accomplish something: looking inside. On the other hand, looking outside ourselves and at other cases is a *distributional* approach to determining uncertainty.

Kahneman-Tversky's Variants of Uncertainty is essential because they attribute uncertainty to *internal* and *external* causes based upon psychological foundations rather than a degree of belief idealized by philosophy, statistics, and decision theory [KT81]. Although this dichotomy is simple and identifies representative attributions of uncertainty, it is also too general and too simple to be useful when identifying specific and critical sources leading to uncertainty.

Throughout the Variants of Uncertainty description, Kahneman-Tversky refer to ignorance several times as a cause for uncertainty. I use their reference and Smithson's Taxonomy of Ignorance as evidence to my contention that ignorance is a major contributor to uncertainty and is appropriately selected as one of the two fundamental classifications for the causes of uncertainty. Their external uncertainty can not be directly related any particular term, at the moment. Through Sections 2.2.4.4 and 2.2.4.5, I identify unreliability as the second fundamental cause for uncertainty. I make my determination based upon some implications of the *external* attributions of uncertainty and ideas inferred from Agosta-Weiss's Sources of Uncertainty (Section 2.2.4.5) and Zimmerman's Causes of Uncertainty (Section 2.2.4.4).

2.2.4.2 Smithson Taxonomy of Ignorance

Smithson establishes a Taxonomy of Ignorance in his book providing a significant amount of detail that I can use to define my own uncertainty taxonomy. In fact, Smithson's taxonomy was the most comprehensive hierarchy I found that included uncertainty. This section provides details that are not repeated in the similar topics of Sections 2.2.4.3 and 2.2.4.5. The ignorance taxonomy and Kahneman-Tversky's reference to ignorance were crucial to my identifying one of the fundamental causes for uncertainty: ignorance.

Smithson explains that uncertainty is a contributor to ignorance, unlike other taxonomies that show ignorance as a source of uncertainty. His defense for uncertainty contributing to ignorance is agreeable as one can appreciate how and why uncertainty can lead to error and

therefore ignorance. Smithson's Taxonomy of Ignorance, shown in Figure 2, and its definitions provide notable information regarding uncertainty; it is revised and used in the final compilation creating my more detailed Taxonomy of Uncertainty. I explain his break down of ignorance then improve it using anomalies found in his own definitions.

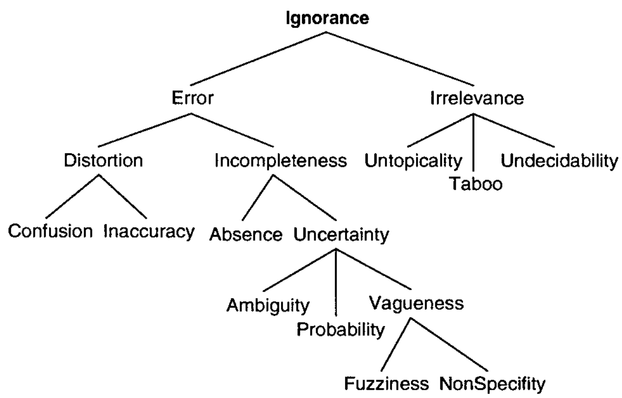


Figure 2. Taxonomy of Ignorance [Smi89]

Ignorance, as defined by Smithson, is the basic lack of knowledge or information for one of two reasons: error or irrelevance. He explains that the commonsense distinction between ignoring and being ignorant is the basis of ignorance. Ignoring refers to the deliberate act of overlooking or avoiding something and is a declaration of *irrelevance*. On the other hand, being ignorant describes unclear or deficient knowledge and is an erroneous cognitive state, hence the term *error*.

2.2.4.2.1 Smithson's Irrelevance

Smithson's *irrelevance* includes three encompassing categories that explain why people might intentionally ignore something: *taboo*, *untopicality*, and *undecidability*. He defines *taboo* as socially inappropriate knowledge and activities. Culture and value systems establish *taboo*, which are the guidelines that identify what people must not know and often not even inquire about. As Smithson puts it: taboo is a "guardian of purity." For instance, cloning and information attack was taboo not too long ago. Topicality refers to the issues of current or local

interest and topical consistency is the intuition that guides ordinary conversation; therefore, *untopicality* indicates those that are not relevant now or here, nor are they discussed. Shop talk, for example, at dinner parties is inappropriate. Smithson explains that *undecidability* occurs when people are unable reach a decision. Sometimes there are issues we are unable to designate true or false because we "consider the problem insoluble" [Sim89] or because its validity or verifiability is not essential or important. Some fantasy and fiction are unverifiable, as are most meaningless thoughts.

I found that Smithson's use of the words undecidability and insoluble are conflicting. Undecidability comes from "undecided' and means indecisive, irresolute, or vacillating, while insoluble means the unfathomable, undecipherable and unsolvable, and describes situations or information that are unknowable. Smithson inadvertently sheds light on another reason for ignorance: the unknown. Unknowable information is distinct, but not incorporated into his taxonomy.

2.2.4.2.2 Smithson's Error

Smithson's categorization of *error*, on the other hand, is a more understandable albeit simple. Simply put, *incomplete* or *distorted* views, information, and processes cause *error*. Smithson indicates that *confusion* and *inaccuracy* cause *distortion*. He explains that *inaccuracy* results in a degree of distortion, while *confusion* indicates mistaken substitution. He then divides *incompleteness* into *absence* and *uncertainty*. *Absence* is when information is simply missing and is the state commonly associated with ignorance. I continue to discuss and extend error and absence in Section 2.2.4.3.

Smithson also explains that uncertainty contributes to incompleteness because "the specificity of the issue" can not be achieved due to *ambiguous*, *vague*, or *probabilistic* situations or information [Smi89]. Although I disagree with his placement of uncertainty, I agree with his basic definitions: *ambiguous* means there is more than one interpretation and *probability* involves

chance. *Vagueness* is defined in terms of *fuzziness* and *nonspecificity*. *Fuzziness* is the lack of clarity and *nonspecifity* is an inexplicit state or information.

2.2.4.2.3 Revised Ignorance

Although the Taxonomy of Ignorance that includes uncertainty as a contributor is well defined, I contend that ignorance more often affects uncertainty rather than the other way around. My proposed uncertainty taxonomy, presented in Section 2.2.5 and Figure 16, reflects ignorance as one of the two fundamental classifications of the causes for uncertainty.

From Smithson's taxonomy, I find that three, rather than two, significant contributors to *ignorance* exist, namely: *error*, *irrelevance*, and the *unknown*. This hierarchy of *ignorance* is an enhancement over Smithson's because of its explicit breakout of the *unknowable*, which Smithson includes within the definition of *undecidability*. Smithson does not explicitly include the issue of the unknown, the unknowable, and the undiscovered except for the confliction I noted in his definition of *undecidability*. Otherwise, in his taxonomy the unknown might fall under *absence* or *uncertainty*. I agree that *error* and *irrelevance* cause *ignorance*, but *unknowable* information is a significant and distinct contributor to *ignorance*. For instance, it is impossible to know the outcome of a game, someone's thoughts, or the exact damage assessment before an attack – they are unknowable. You could argue that taboo and the untopical seem like they could be considered unknowable, but they exist and can be learned. Whereas the unknowable, that should be encapsulated, is that information or knowledge which is not possible to know. The distinction between the *erroneous*, *irrelevant*, and *unknowable* form the Revised Hierarchy of Ignorance shown in Figure 3 and used later in my Taxonomy of Uncertainty. Error is truncated here because it is revisited and extended in Section 2.2.4.3.

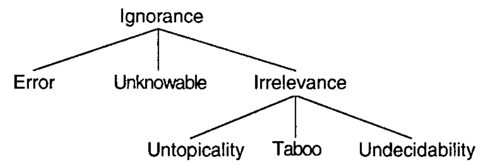


Figure 3. Revised Hierarchy of Ignorance

2.2.4.3 Giarratano-Riley Types of Errors

Giarratano and Riley categorize types of error [GR94] providing more specific details (see Figure 4) than Smithson does in his Taxonomy of Ignorance (see Figure 2). They engulf and extend the essence of error described in the Taxonomy of Ignorance by providing a detailed hierarchy of the types of errors that occurred in or relevant to expert systems. They explain that these errors contribute to uncertainty, which is in contrast to Smithson's view that uncertainty contributes to error. Their breakdown identifies seven distinct type of error (see Figure 4) versus the two generalizations classified by Smithson (see Figure 2). The Giarratano-Riley seven types of error are *ambiguous*, *incorrect*, *incomplete*, *random*, *measurement*, *systematic*, and *reasoning*. For the most part, the types of error are self-explanatory and do not require much explanation. I reorganize both sources for error into four subdivisions identifying distinct reasons for error that used to create my Taxonomy of Uncertainty in Section 2.2.5. The only comments I have involve incorrectness, human error, reasoning, and measurement.

Giarratano-Riley include a crucial elaboration on *incorrectness* rather than simply attributing it to inaccuracy. *Human error*, *malfunction*, and *false positive* or *negative results*, they explain, cause *incorrectness*. Giarratano-Riley also segment the errors that could be caused by a person. They explain that errors in reasoning are distinct from the simple mistakes that people make, which is why they identify *reasoning* as one of their seven types of errors and *human error* as a sub-type to *incorrectness* (see Figure 4). *Reasoning* is affected by *inductive* and *deductive*

errors. Another focal point of their types of error is *measurement*, which is affected by *precision* and *accuracy*.

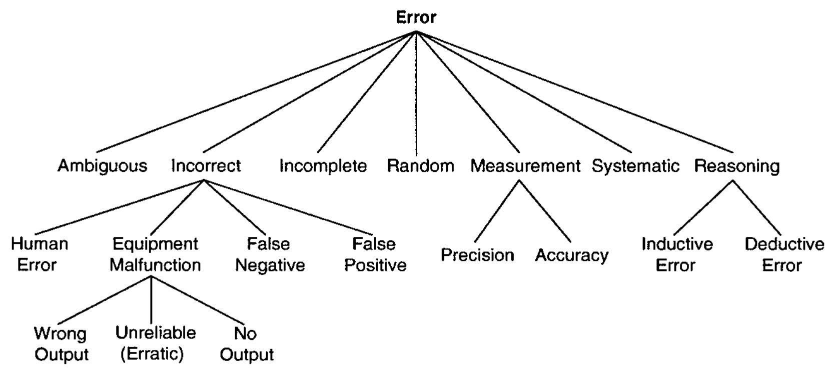


Figure 4. Types of Error [GR94]

I reorganized the Giarratano-Riley Types of Error (see Figure 4) and Smithson's Taxonomy of Ignorance (see Figure 2) to create the breakdown shown in Figure 5. By regrouping the types and terms identified by both sources I was able to create more distinct and appropriate categories based upon causal similarities. *Reasoning*, *measurement*, *incompleteness*, *incorrectness*, and *distortion* are appropriate subdivisions that cause *error* and encapsulate all or some revision of the elements identified by Giarratano-Riley and Smithson.

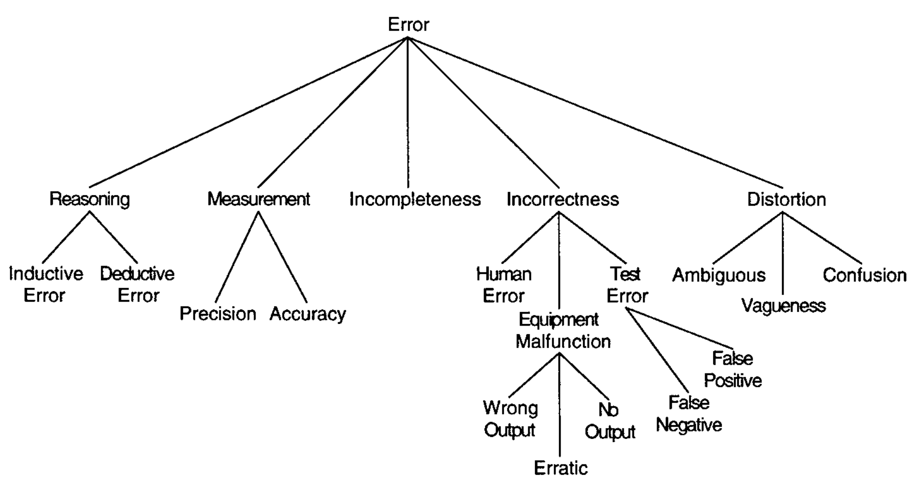


Figure 5. Causes of Error

2.2.4.3.1 Reorganizing Error

Figure 5 shows some clear differences to the previously mentioned ideas of error, particularly the reorganization and the category *distortion* taken from Smithson.

First, I tried to capture the reason for basic errors that occur everywhere under *incorrectness*. This also reduces the Giarratano-Riley Types of Error breakdown into four common causes: *reasoning*, *incompleteness*, *incorrectness*, and *distortion.* Second, I use Smithson's *confusion*, *ambiguity*, and *vagueness* that cause misrepresentation to create the *distortion* category. Third, Smithson's inaccuracy is captured under Giarratano-Riley's *measurement*. Fourth, The Giarratano-Riley *incomplete* and Smithson *incompleteness* categories are refined into a singular *incompleteness* that identifies errors caused by incomplete processes, details, and more. Fifth, I use Smithson's *absence*, and Giarratano-Riley's *random* and *systematic* categories as well as their explanation for "incomplete" to define the category *omission.* I also determined that *omission* contributes directly to *ignorance* extending its hierarchy as shown by Figure 6.

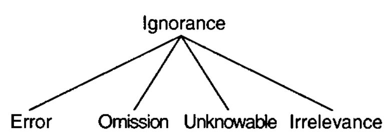


Figure 6. Extending the Ignorance Hierarchy

I chose to locate *omission* under *ignorance* because omitted information results in a lack of information more so than it leads to error. Unlike irrelevant and unknowable things, omitted details and information are distinct since they are attainable, but missing. *Omission* or missing things occur from intentional as well as random actions, which is why *systematic* and *random* are placed under *omission* as shown in Figure 10. *Systematic* omission includes intentional absence and removal of information further indicating that omission clearly refers to things that can be learned rather than those that are unknowable or irrelevant. This is similar to *irrelevance* in the sense that it can be learned, but significantly different since factors determining *irrelevance* are

internal to our person or society while *omission* is a factor employed to preclude access to something. In addition, since I define *omission* as a removal of information it would not include the unknowable or be included under *unknowable*.

The categories omission, systematic, and random were removed from error because I felt they contribute more to ignorance (see Figure 10) than error. On the other hand, using information that is unknown or missing for any reason in a process can lead to error, which is why someone might think that the categories *omission* and *unknowable* should be placed under *error*. A formula or process that attempts to include unknown or missing details would itself be erroneous by the logic or *reasoning* (*inductive* or *deductive*) that created the formulation, not because of the data or the lack of that data.

2.2.4.4 Zimmerman Causes of Uncertainty

Zimmerman's paper on fuzzy decision support systems includes a breakdown of several causes of uncertainty: the lack of information, abundance of information (complexity), conflicting evidence, ambiguity, measurement, and belief [Zim98]. The notions of belief, complexity, and confliction are unlike other causes or terms presented in Sections 2.2.4.2 and 2.2.4.5. These terms do not reflect ignorance and become part of another reason for uncertainty. Zimmerman' terminology actually helps identify the second fundamental classification for the causes leading to uncertainty: *unreliability*.

Belief is a subjective consideration of a situation or information and raises an issue of credibility. The cause "conflicting evidence" also reflects an issue of credibility, particularly the credibility of the source of each piece of conflicting evidence. These two factors led to the identification of the category *credibility*. Zimmerman oversimplifies several other reasons for disbelief that I omit because they can be encapsulated under inconsistency and contradiction. *Inconsistency* and *contradiction* tend to mean the same thing and are placed under *credibility* as a

single item. Complexity, on the other hand, does not fall under either of the categories identified so far and is revisited in Section 2.2.4.5. However, this forced the identification of a suitable category that could include credibility and some notion of complexity. The Agosta-Weiss Sources of Uncertainty described in Section 2.2.4.5 helped to identify the category *unreliability,* which *credibility* and complexity suitably fall under.

2.2.4.5 Agosta-Weiss Sources of Uncertainty

Agosta and Weiss categorized the Sources of Uncertainty in Figure 7 as the fundamental challenge of medical diagnosis and troubleshooting [AW99]. Although their break down concentrates on problems of the medical domain, they provide insight that helped identify *unreliability* as a suitable term that corresponds to Kahnemann-Tversky's external attribution of uncertainty. The Agosta-Weiss Sources of Uncertainty are founded on three major causes for uncertainty in medical diagnosis: *incomplete data*, *noisy data*, and *weak discriminators*.

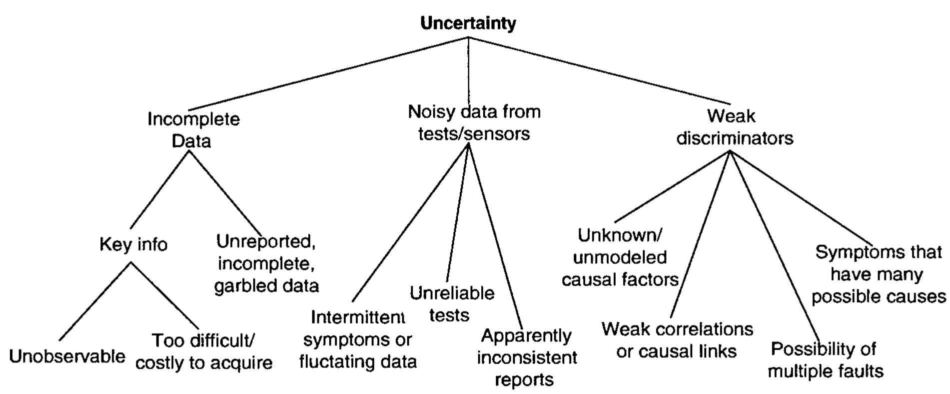


Figure 7. Sources of Uncertainty [AW99]

In their hierarchy, *incomplete data* is caused by *key information that is unobservable* or *too difficult or costly to acquire* as well as *unreported, incomplete, or garbled data.* Unobservable information can be encapsulated by the *unknowable* category in *ignorance.* Information that is too costly or difficult to acquire is representable by *omission* since it is

achievable, but intentionally avoided. The other reasons *unreported, incomplete, or garbled data* are already identified under *ignorance.*

The Agosta-Weiss category of *noisy data* is represented by *distortion* that causes *error*as shown in Figure 5. They explain that *intermittent symptoms*, *unreliable tests*, and *inconsistent reports* cause noisy data. Their last category, *weak discriminators* is not explicitly identified in previous explanations, but can be captured under *incorrectness* or *measurement.* Its specific causes *unknown or unmodeled causal factors* are represented by the *unknowable* and *omission,* while *weak correlation or causal links* can be identified with *vagueness* and *ambiguity* under *distortion.* The *possibility of multiple faults* and *symptoms that have many possible causes* are also representable by *ambiguity.*

Although most of the Agosta-Weiss sources for uncertainty were classifiable under previously defined terms, their issues pointed out two other major causes leading to *unreliability*: *limitations* and *acquisition and exposition.* *Unreliability* is defined in terms of *credibility, limitations*, and *acquisition and exposition* as shown in Figure 8.

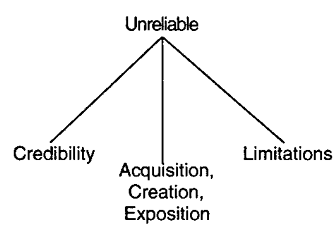


Figure 8. Causes for Unreliability

I determined that the complexity mentioned by Zimmerman (see Section 2.2.4.4), the problems caused by sensors, and the difficulty of diagnosis reflect issues related to *limitations. Limitations* identify the bounds within which something or someone operates, which is explained further in Section 2.2.5.2.2. I also realized that the other problem they were having was relative to the process of gaining and reporting information. In their case, their patient interviews, reports, and tests posed a problem because of the problems that occurred during these events.

With this in mind, I created the category *acquisition and exposition* to encapsulate causes for uncertainty that occur as data is being acquired or reported. This was refined to include *creation* to include weaknesses that introduced during manufacturing processes. Section 2.2.5 continues this break down.

2.2.5 Taxonomy of Uncertainty

The following section assembles the different versions and sources of uncertainty presented in Sections 2.2.4.1 through 2.2.4.5 with other details to create my Taxonomy of Uncertainty. Figure 9 shows the basis for the taxonomy. I organized the causes for *uncertainty* into two fundamental classifications: *ignorance* and *unreliability*, each of which is subdivided into several sources. As previously mentioned, I established this taxonomy of causes for uncertainty to help define a technique for visualizing uncertainty. In the process, I created a taxonomy that is more specific, thorough, and accurate than any other classification I found. This less ambiguous taxonomy is useful for dichotomizing one cause for concern or uncertainty from another with increased precision.

The taxonomy's explicit organization also provides a means for standardizing the identification of uncertainty, particularly in fields and environments where uncertainty is a common feature. Intelligence analysts and decision-makers, a focal point of this thesis, commonly deal with uncertainty, a well-defined Taxonomy of Uncertainty could facilitate the inclusion of uncertainty in their decisions and their decision support systems. The taxonomy could improve understanding of certain issues as well as explain underlying reasons for others. The taxonomy can also aid my selection of methods and symbols for visualizing uncertainty.

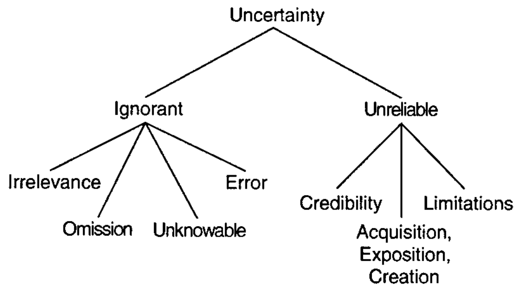


Figure 9. Basis for Uncertainty Taxonomy

Finding a classification of uncertainty was not as effortless as expected; there were plenty of vague and assuming discussions of and including uncertainty but very few analytical dissections of uncertainty. Although there are some similarities between the previously discussed categorizations, neither itemization was in complete concurrence nor did they match the internal and external sources of uncertainty identified by Kahneman-Tversky. Consequently, it was necessary to produce another taxonomy that incorporated the concepts presented by respected researchers.

Determining the most judicious and basic reasons for uncertainty was the key to establishing this Taxonomy of Uncertainty. Most research supports the idea that uncertainty is attributable to internal and external factors [Cle96, Dah96, KT82, Smi98], which I used to guide the selection of phrases that capture both issues. I chose *ignorance* and *unreliability* as the fundamental classifications of the causes for uncertainty during my examination of supporting material.

The two most general categories leading to *uncertainty* are the lack of knowledge and fallibility of sources, information, and processes: *ignorance* and *unreliability*. The classification *ignorance* stands out through Kahneman-Tversky's Variants of Uncertainty and Smithson's Taxonomy of Ignorance. The *unreliability* classification was realized in a pattern I noticed in the material that indicated imperfection or unreliability. These classifications are justified in four ways. First, via Kahneman-Tversky who make several references to ignorance as a significant

cause for uncertainty as they explain the Variants of Uncertainty in Section 2.2.4.1 and shown in Figure 1. Secondly, inverting Smithson's Taxonomy of Ignorance (see Figure 2) indicates that ignorance contributes to uncertainty, which is reasonable by commonsense. Third, throughout many of the uncertainty models or their decomposition, the most general category engulfing issues is reliability, ideally the reliability of the processes and information. Finally, reliability is identified as a problem in the Agosta-Weiss Sources of Uncertainty discussed in 2.2.4.5 and shown in Figure 7.

In terms of relating to the internal and external sources of uncertainty, ignorance tends to reflect internal issues while external events and elements cause unreliability. The relationship to internal and external factors fades as ignorance and unreliability expand in meaning. The loss of correlation mainly occurs as each classification encompasses causes leading to uncertainty that involve internal as well as external factors. For instance, things, information, and data cannot be ignorant, but the source can be erroneous effectively resulting in missing knowledge.

The following sections compile *ignorance* and *unreliability* to complete the hierarchical classification of the causes for *uncertainty*. The discussion explaining the *ignorance* classification is short since I cover most of it throughout the previous sections: Sections 2.2.4.1 through 2.2.4.5. On the other hand, *unreliability* was not fully expanded on and is the focus of most of this section.

2.2.5.1 Ignorance

Ignorance is specified and compiled from previous sections creating a more complete hierarchy of ignorance. Combining the Extended the Ignorance Hierarchy shown in Figure 6 and the refined Causes of Error in Figure 5 creates the comprehensive categorization of *ignorance* shown in Figure 10. From this point on, ignorance is defined as the lack of knowledge or information due to one or more reasons identified within the four causes for ignorance: *error*,

unknowable, *omission*, and *irrelevance*. The complete hierarchy of uncertainty, including ignorance, is in Appendix A.

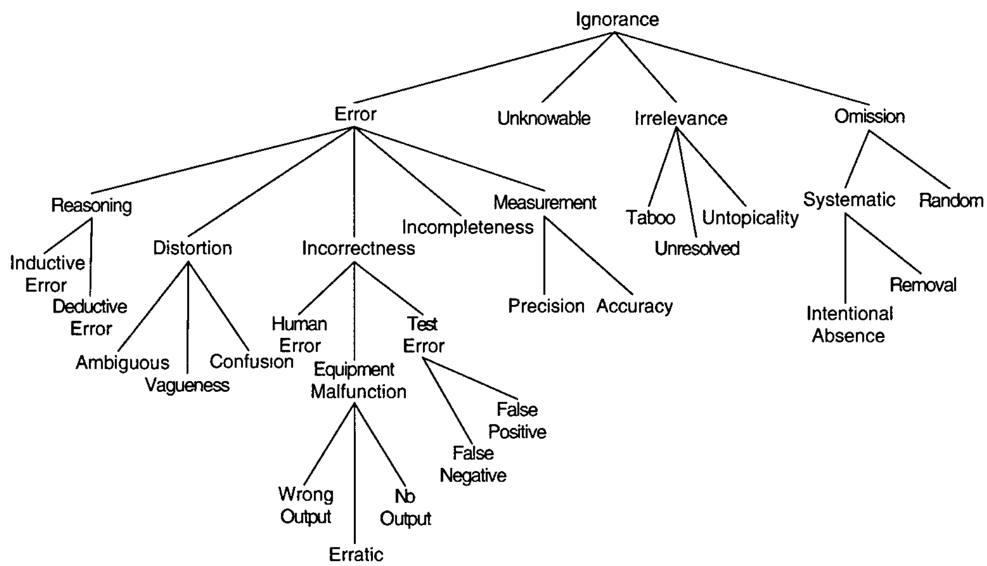


Figure 10. Final Categorization of Ignorance

To see the classification of ignorance "work" suppose an analyst is gathering information from open sources about a new Soviet tank. In addition, suppose most of the data provides details about the tank's speed, communication capabilities, and defenses, yet nothing regarding the main gun. The resulting knowledge about the tank is missing data, which leads to ignorance. The missing details could have been caused by the *omission* of details via censorship or because they were *irrelevant* to the source. The resulting collection of information includes knowledge of a tank, but has uncertainty relating to the main gun. It is not uncertain because of miscalculations, old data, taboo, or equipment malfunction, but because some data was absent.

Unfortunately, some strategies for adding information to a DSS preclude adding the tank until all details are certain. From personal experience I can attest to the fact that strict systems are not always better. Systems following difficult guidelines such as requiring data to be completely accurate before allowing it to be added to the knowledge base can frustrate users and drive them to simplifying data, including false data, and completely omitting data. I suspect that DSS are

subject to similar issues, especially when considering the human factor. This further supports my contention that we should include and visually express uncertainty in DSS.

Continuing the previous example about the tank, a person faced with the issue of not being able to add details about it to the system simply because one piece of information is missing might try to find a way around the constraint. Human commonsense will reason that the tank most likely has main gun that is typical for similar tanks, say 105mm. Furthermore, considering that it could be deadly not to include the tank or its information, the person may choose to augment the information with their own inferences to protect others. Unfortunately, by doing this the analyst extends the *ignorance* to the DSS data in terms of *inaccuracy* and compromising his *credibility* and further damaging the knowledge base. The details about the tank and the lack of information could be added to the system without compromising the analyst or the data if there was an appropriate method for identifying and expressing the uncertainty.

I speculate that in many systems the data about the tank would probably be added with some potentially incorrect data. In some sense, it is better to be overly cautious than careless. This simple scenario demonstrates the application of the taxonomy and value added for including and expressing uncertainty. Furthermore, it shows that we should not be surprised if our own data is also wrong for reasons other than the reporting source, system design and human interaction insert additional uncertainties.

2.2.5.2 Unreliability

Unreliability is the other fundamental subdivision contributing to *uncertainty*; it contains many reasons and factors to question the reliability of information, data and processes that are referenced. Unlike the preceding section about *ignorance*, this next section is extensive due to the relative lack of previous information regarding unreliability.

I identify three categorizations of the reasons for *unreliability* in Section 2.2.4.5, which are shown in Figure 11. I determined that *reliability* associated with information and objects or the lack thereof depends directly upon the *credibility* of the source, *acquisition*, *creation*, or *exposition* process, and *limitations* affecting sources and processes alike. Unreliable information is often one of the greatest factors affecting information gathering. Intelligence analysts regularly contend with issues of credibility as well as those occurring during the process of acquiring and reporting information. Extending *unreliability* becomes increasingly important and difficult.

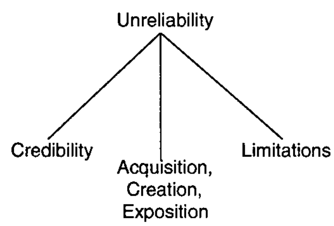


Figure 11. The Basis of Unreliability

The numerous fields and areas contributing to reliability provide many opportunities for overlap and repetition making the distinctive breakdown of *unreliability* difficult. For instance, a major reason to question the reliability of information is its accuracy, which is defined under *ignorance*, but might be identified under *unreliability* if there were details as to why it was inaccurate. On the other hand, we could identify both aspects of the uncertainty: the inaccuracy of the information and why it is inaccurate.

You could argue that accuracy is better suited in the category *limitations* because, for instance, the accuracy of sonar is only as good as its equipment and algorithms. I categorized my taxonomy so that a well-stated cause for uncertainty will fall into one of the respective areas. As the previous example demonstrates, multiple interpretations and causes for uncertainty are identifiable when more information is known. The overall intent was to provide a breakdown representative of the real world and free from repetition, overlap, and similarities; not to preclude diverse interpretations and insight into information. If a situation occurs where there is or might

be confusion or overlap, simply pick the most suitable area and avoid duplication. We can establish heuristics to resolve potential overlap such that overlapping details are placed in the more important categories, such as credibility, if these have been identified.

2.2.5.2.1 Credibility

Credibility or the lack of it tends to be associated with information, processes, people, abilities, and events; it is not commonly associated with objects or things. For instance, we do not question the credibility of a cracked cup, we do however, challenge its ability to retain fluids. I defined *credibility* by considering those effects that affect credibility or cause it to be questioned all the while avoiding the category of ignorance. Figure 12 shows the *credibility* breakdown that is clarified in the following section. Various sources and brainstorming led to this hierarchy. The following list identifies six reasons to question *credibility*: *probabilistic*, *deterministic and hierarchically*, *influence*, *inconsistency and contradiction*, *misuse*, and *tainted.*

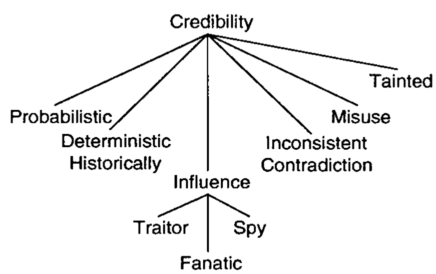


Figure 12. Credibility

Data that is *tainted* for some reason affects the *credibility* of a report or information source. Another way to considered information *tainted* is to base the uncertainty on the messenger or source of the information who may be untrustworthy, in which case the data could become *tainted* by simply going through his custody. I have identified several possible interpretations in some of my examples to point out that as interpretations increase so do the potential inferences. The differences should be acceptable as long as the uncertainty is identified and accounted for accurately or as accurately as it can be discerned. On the other hand, I would

not suggest making wild interpretations to try to come up with multiple variations. Insight to uncertainty should come naturally from the information provided, otherwise you may introduce your own uncertainty through your inferences. Perception and understanding also affect the uncertainty that is identified; what one person sees as an issue of *credibility* could be identified by another person as an issue tied to *ignorance*. Is either person more correct? Neither person is necessarily wrong; they see things differently and are correct in that sense. Both identify some reason for uncertainty and both maybe valid. The opportunity for multiple interpretations is not negative and does not reflect a poorly designed taxonomy; it reflects the fact that people can perceive different information from the same data.

Continuing with the classification of *credibility*, I explained in Section 2.2.4.4 that *inconsistent* and *contradicting* information are simple indicators that should cause someone to question the credibility of a source and the information itself. *Inconsistency* is the occurrence of more than one plausible solution, while *contradiction* is when the solutions or information are conflicting or opposing. These are grouped together because they indicate there are two or more different indicators or pieces of information that cause you to question which one is more accurate. Furthermore, if the details came from different sources, there may be other implications that reflect on the *credibility* of the sources.

Misuse of data and resources are indicators that should raise questions about the integrity of the data as well as the source. The media, for instance, often misuses statistics. For example, handgun opponents often misrepresent the number of deaths due to firearms (to their benefit) without identifying the statistics about those that were accidental or otherwise non-criminal. *Misuse* also occurs when information is taken out of context, as the media often does. The credibility of a source and his information should be an issue when we know that other groups, factions, or events can *influence* him. For example, we should be leery of the spies the US has converted into double agents who feed the opponents half-truths. Traitors, prisoners, fanatics,

and addicts lack *credibility* as they will do and say almost anything to get what they want; they also tend to have uncontrollable fears and problems that could influence them.

Probabilistic references are associated with chance, while *determinism* and *history* have some predictability. These three areas capture the processes, events, and data that have some pattern or trend that indicates why *credibility* should questioned. For example, a process that generates information with only 30% accuracy has limited *credibility*. From another perspective, if a person has to blindly pick a data point from several choices, then the *credibility* of the selection is based upon the *probability* that he selected the correct one. On the other hand, *deterministic* and *historical* trends can also be used to indicate the undesirable lack of *credibility*. For instance, a reporter or source that routinely provides inaccurate data will be barely *credible*, as new information is acquired from them one can expect it to be inaccurate. A deterministic example might be the source that exaggerates when bragging but discredits under normal conditions. The key to determining *reliability* is to identify the source or situation surrounding the event or information, which requires additional information and knowledge.

2.2.5.2.2 Limitations

This section on *limitations* covers Figure 13 and the category where constraints and hindrances affect the reliability of the source, process, data, information, and object. Unlike credibility, limitations are associated with objects, abilities, and information alike. English is an example of a language with limitations imposed by its ambiguities as well as its simplicity. In English, like other languages, meaning depends upon usage (syntax and context). For instance, the word 'take' can be a verb meaning to retrieve something, or noun referring to your winnings or thoughts. We can also inadvertently lose information in translation. Clearly, 6000 actively spoken languages in the world today [Lea99] result in countless limitations lost in translation. Other *linguistic* limitations include the lack of specificity in words, particularly those used for fuzzy sets and fuzzy logic. Fuzzy sets duplicate the vagueness of natural languages.

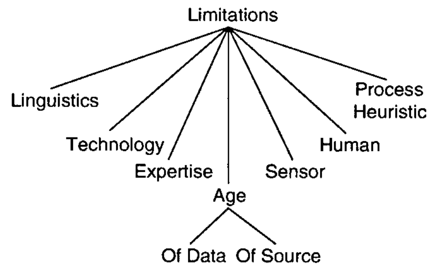


Figure 13. Limitations

Sensors have limits too. When data is acquired using tools at or beyond their limits, we need to be wary of their readings, data, and results. Most sensors are designed to work and sense in one or a few domains and will be incapable of providing any other data. We might question the credibility of the data and the reporting source when data is provided outside a realm of a sensor. Like sensors, *human* receptors and senses have limits. Our vision, hearing, speed, strength, and cognitive load are among some of our *limitations*. Obviously, we are not likely to observe a man walking on the moon with our naked eyes. On the other hand, there are many things some people can do that others cannot. The incredible feats should not be disbelieved, but might have other associated uncertainty and will raise concern until the feat is proven or accepted.

Technology can be a limiting factor in two ways: when it is too expensive to build or achieve and when we do not have the ability to build or accomplish a task or feat. Nuclear energy through fission is possible; however, it is not practical, as we do not have the capability to trap and control the release of energy. *Expertise* is a little like technology in the sense that we can have some, but we cannot have it all. This area would be used to capture and record those conditions or things that we are somewhat unsure about because the source is limited in its expertise, technology, or capabilities.

Age as a *limitation* can be attributed to data as well as the source. The importance of *data's* age depends upon its application and is a *limiting* factor when the quality of the data is tied to its age. A two-week old weather report has no value whereas a forecast of tomorrow's weather from two weeks ago may be imperfect, but has more value than no report at all. The *age of* a

source can also be a limitation. For instance, younger people see more color than the elderly do. In addition, perspectives can be different: a child's view of one scene will be different from that of an adult and possibly that of an elderly person. Some people may not even understand what they are witnessing and are unable to report the correct information.

Automation and *processes* are also subject to *limitations* and mistakes, particularly if they incorporate various human-like *heuristics*. We incorporate various *heuristics* as part of our natural reasoning and as Section 2.1.2 indicates, our *heuristics* are imperfect and affected by bias, which can lead to incorrect reasoning and results. Although bias is not directly represented in my taxonomy, it occurs for many reasons without being discussed as a bias. As explained in Section 2.1.2, we are all affected by a variety of biases that can vary by situation and surroundings. Bias is introduced with age, stress, desire, irrelevance, expertise, and more. Therefore, bias is present throughout this taxonomy without directly calling attention to it.

2.2.5.2.3 Acquisition, Creation, and Exposition

The processes of *acquiring* and *reporting* information as well as the processes for *creating* objects can inject impurities making them weak or prone to damage. Figure 14 identifies several causes for problems and negative affects to the processes of *acquisition*, *exposition*, and *creation*. *Interference*, *denial and deception*, *transformations*, *stress*, *environment conditions*, *decoding*, and *mishandling* affect these processes.

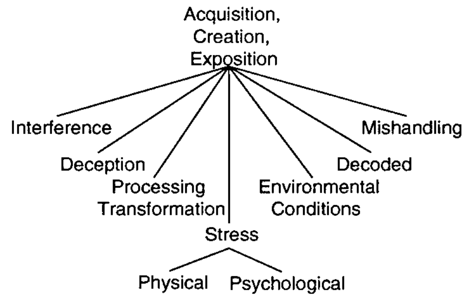


Figure 14. Acquisition, Creation, and Exposition

Having knowledge of the active *denial and deception* practices of any entity is an obvious inspiration for uncertainty. Information gathered from a source that practices *deception* might be completely or partially incorrect or misleading. Information that was encrypted and forcefully *decoded* should also be considered with some degree of uncertainty. It is possible that the information or cryptographic algorithm was intentionally weak as a strategy for *deception.* The uncertainty decreases as confidence in the algorithm increases and the compromise remains undetected, much like World War II and the German encryption algorithm cracked by a Polish mathematician in 1933 and used by the allies. The cracked algorithm allowed the allies to decipher the German communications for the entire war.

Another event that can cause problems is simple *mishandling* of the information and data. *Mishandling* is introduced when the information or media transporting the information is affected, such as burnt records and magnetic interference that wipes out portions of a magnetic medium. The information is blurred, damaged, or removed due to *mishandling*; you would be unsure about its quality and completeness when this happens.

Other factors that affect the process of gathering and reporting information include *environmental conditions* and *interference.* Solar flares interfere with communications and if an untimely flare wipes out a segment of data then some information will be missing and the clarity of the information reduced. *Processing* and *transformation* could be problematic, particularly

when involving equipment, algorithms, and methods that poorly or incorrectly work with or use the information.

The final area that can cause one to be wary about data or information is *stress*. People tend to behave differently and unpredictably under *stress*; perceptions and views of the world can even change. I discretely identified *stress* as a cause for uncertainty in the *acquisition*, *creation*, and *reporting* process because it changes our perceptions and responses. Figure 15 identifies several causes for *stress*, which I compiled from several sources as well as through self-reflection; Hockenbury, [Hoc97] and Reed [Ree92] are two sources.

I segmented *stress* into two distinct areas: *physical* and *psychological,* portraying many, if not most, factors contributing to *stress* as shown in Figure 15. A person reporting information under stress might exaggerate or omit something as well as underscore or forget other data. Victims of crime have a hard time of accurately reporting details; investigators can employ various strategies to elicit 35-60 percent more detail [Ree92]. Various uncertainties can and should be associated with information and data acquired and reported under duress and *stress*.

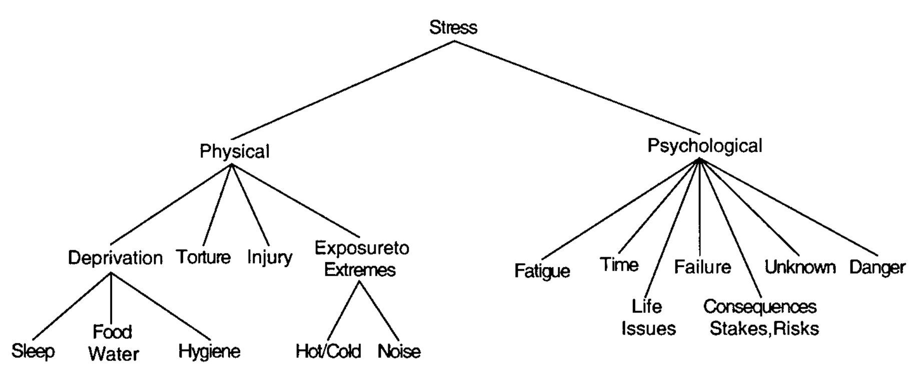


Figure 15. Sources of Stress

The previous sections provide hierarchical categorization for each contributor that results in *unreliable* information or data. *Limitations*, *credibility*, and *acquisition* or *exposition*s are causes for uncertainty leading to the classification *unreliability*, which I have identified as the

second fundamental classification of the causes for *uncertainty*. I branded *ignorance* as the other fundamental classification of the causes for *uncertainty*.

2.2.5.3 The Completed Taxonomy of Uncertainty

At the start of this Section, 2.2.5, I indicated that the two fundamental classifications of the causes for *uncertainty* were *unreliability* and the *ignorance*. Combining these two classifications creates the Taxonomy of Uncertainty shown in Figure 16. This taxonomy shows the division of uncertainty into the two fundamental classifications (unreliability and ignorance), each with their distinct types ((error, irrelevance, unknowable, omission), (credibility, acquisition/creation/exposition, and limitations)), and numerous reasons, sources, and causes for each. Appendix A contains a fully expanded version of the Taxonomy of Uncertainty with each term and classification I have presented.

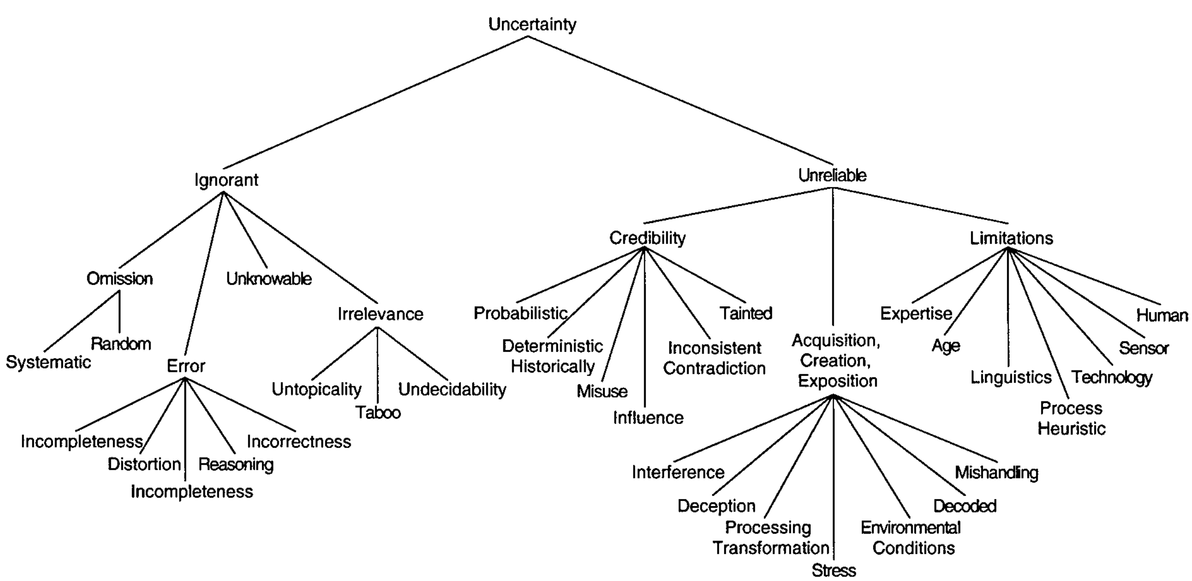


Figure 16. The Taxonomy of Uncertainty

The Taxonomy of Uncertainty facilitates the characterization and categorization of conditions, events, and sources associated with and resulting in uncertainty. This taxonomy provides a wide-ranging categorization of the causes for uncertainty that should encapsulate and

include virtually every reason for uncertainty. I designated distinct categories that can have many meanings and be comprised of numerous problems, issues, and reasons. These categories were identified with this intent because it would be impractical to attempt to include every single reason for uncertainty.

Most of the *ignorance* classification is founded on the research and taxonomies identified by various experts, some of which are presented in Sections 2.2.4.1 through 2.2.4.5. *Unreliability*, in contrast, has few specified references, but is founded on many details discussed throughout material I explored. Unfortunately, none of the material specifically identified and explained a hierarchy relating directly to *unreliability* or any of its three causes; I derived these from the details the resources provided.

In terms of applying my taxonomy, I realize it is possible to have many interpretations of a given scenario; however, each interpretation will be reasonably similar or distinctly identify another reason for uncertainty. For this reason, I believe that there is nothing terribly wrong with sorting or classifying data differently as long as the information and reason for the uncertainty are retained. The choice for identifying uncertainty can depend upon the causes for concern, the quality and quantity of information available, the type of processing that influenced the information, and the linguistic measures used by the provider, observer, and analyst [Zim98].

2.2.6 Identifying and Expressing Uncertainty

Decision-making can be very difficult to do, particularly in military environments and when involving lives. The complexity, uncertainty, undesirable results, and different perspectives present in every situation make decision-making difficult. One aspect of my research was to find a technique for expressing uncertainty that I expected would provide the user with a different perspective, which I theorized would be beneficial. A different perspective could be advantageous. The focus of my work was to enhance the information used by the analyst and

decision-maker by including uncertainty. Uncertainty would have to be included in a way that was intuitive and caused the fewest problems for the user. One of the problems I wanted to avoid was the challenge of handling uncertainty, in particular having to read a variety of numbers in order to understand the uncertainty that was being identified. With this in mind, I sought a technique for handling uncertainty that would also be used for identifying and expressing the uncertainty.

I explored many different methods for handling and expressing uncertainty including classical approaches and finally settled on a more novel concept: fuzzy logic. Probability theory, Bayes theorem, Bayesian networks, certainty theory, nonmonotonic logic, and evidential reasoning were among the approaches that I did not feel were appropriate for use in a DSS where uncertainty would be included to extend the information available. Fuzzy logic was most suitable; its use of natural language rather than numbers does not imply perfection and could be saved for interpretation by the user and other system processes.

2.2.6.1 Fuzzy Logic

I determined that fuzzy logic provided the best solution identifying, expressing, and handling uncertainty that is present in a system commonly involving people, language, imprecision, and differences of opinion. Its use of natural language fits neatly into the intelligence analysts' domain where common words make up a significant portion of the data. Natural language, words, and fuzzy terms can encapsulate the data as well as the uncertainty. I explain in Section 3.2 why I believe the retention of these descriptions is critical, mainly as a means of facilitating dynamic analytical evaluation by automation and the end user.

As previously noted, fuzzy logic differs from classical logic in that values are no longer quantitative or numeric. The fuzzy approach uses words and terminology of the human language to capture meaning and membership functions or intervals to define values they represent.

Although composed of natural language, fuzzy logic is a superset of Boolean logic that is extended to handle the concept of partial truth composed of values between completely true and completely false or [0, 1] as with other approaches. Zadeh defined fuzzy logic as a means to model the uncertainty and vagueness of natural language [MF93]. Consider the example of a man's height, it can be generalized to three sizes: tall, medium, and short. Figure 17 shows a graph representing the values most people assign to those three sizes. Notice that each value overlaps and covers more than one size. The overlap represents the different opinions of others; for instance, very few people consider the 6'5" to be short while everyone considers 7' tall. In addition, notice that the term "tall" is also vague in the sense that you are unclear about the exact value that it refers to; in essence, "tall" captures both the data and uncertainty.

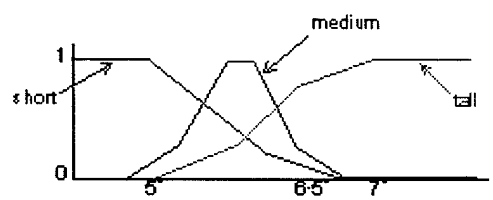


Figure 17. Fuzzy Graph of Tall

Zadeh denotes that rather than "regarding fuzzy theory as a single theory, we should regard the process of 'fuzzification' as a methodology to generalize any specific theory from a discrete to a continuous (fuzzy) form" [CMU93]. The extension principle of fuzzy logic makes this possible by providing techniques for converting from the discrete system to the fuzzy form. For example, it is possible to create a Bayesian network using fuzzy terms instead of probabilities.

One of the greatest challenges of fuzzy logic is the process for creating the fuzzy models. The process is "fuzzification" and involves decomposing system input and/or output into fuzzy sets. The designer must create or identify the membership functions and models that describe the vocabulary as well as the methodology for handling the terms before applying fuzzy logic to a system. Consider, for example, a home thermostat using fuzzy logic. Its terms might include

very hot, *hot*, *warm*, *cool*, *cold*, *very cold*, and *off*. These terms can represent slightly different values for each person. We establish our assessment of each term based upon our experiences of different environments and temperatures; e.g. 65° F is *cold* to me. System designers would determine the fuzzy set values and their ranges through extensive surveys (or guesswork) to define membership functions that reflect the majority of the population or customer base. In terms of the thermostat, the intervals (or membership functions) for *warm* could be 70-79 degrees while *cool* was 67-72 degrees. The opposite process, defuzzification, converts qualitative terms and descriptions into a single value that best represents the fuzzy set or membership function.

2.2.6.1.1 Naturally Descriptive

Through its use of naturally descriptive quantities, fuzzy logic is intuitive to human-use and decision-making. Individual meanings for various words may differ (e.g. very likely is about 80% to some and 60% to others), but their approximations are similar. On the other hand, when we use or are given specific values (e.g. 80%) we are provided with potentially imperfect information and an implied precision about the uncertainty. Unlike probability-based approaches, the natural language of fuzzy logic avoids the implied precision of numbers. As previously mentioned, numeric values can be converted (fuzzified) into representative fuzzy words and vice versa where fuzzy terms are defuzzified into numbers.

Table 2 provides some natural language terms that express quantities, which makes them fuzzy terms. Table 2 also identifies some hedges and modifiers, these terms modify the value represented in the fuzzy sets by increasing, decreasing, squeezing, stretching, or shifting the data interval and the uncertainty identified by the membership functions.

Table 2. Some Fuzzy Terms

Quantifying Words	Occasionally	Often, likely	Almost always	Once in a while	Sometimes
	Unlikely	Possible	Probable	Frequently	Infrequently
	Usually	Unusually	Improbable	Always	Very often
	Never	Almost never	Seldom, rarely	Very seldom	Often
Hedges or **Modifiers**	About	Around	Above	Positive	Below
	Vicinity	Generally	Close	Not	Somewhat
	Very	Extremely	Slightly	After	Before

Fuzzy logic is an approach to handling uncertainty and provides a variety of operations for manipulating the terms and represented values. Table 3 identifies some of these fuzzy set operations. Further details of fuzzy logic and these operations is beyond the scope of this work.

Table 3. Fuzzy Set Operations

Union	Intersection	Complement	Intensify	Mean-and
Product-and	Product-or	Threshold-not	Cosine-not	Mean-or
Bounded-and	Bounded-or	Zadeh-and	Zadeh-or	Zadeh-not

2.2.6.1.2 Why Fuzzy Logic?

I found that fuzzy logic was the most appropriate way to represent and express uncertainty primarily because of its use of natural language and fuzzy sets that reflect human communication. In Chapter 3, I recommended the use of fuzzy logic as part of the approach to accomplishing my goal of enhancing information used by analysts and decision-makers in DSS. Using fuzzy logic is one part of identifying the uncertainty that is included in a DSS; the second is expressing it visually. In this case, it is also possible to use the problem domain of intelligence gathering and its experts as a means to justify the selection of fuzzy logic for handling uncertainty: words and documents are a source of their data.

Intelligence gathering depends on people to both reveal information and translate it into usable knowledge. It is sensible to record the information they provide in the form that is most natural for them. For most people, the natural form is our natural language, which includes letters, symbols, words, and numbers. Furthermore, we tend to communicate in words and sentences that are imprecise rather than numbers and formulas, which indicates that analysts and

intelligence gathering techniques must accommodate or translate these words. Translation itself can introduce uncertainty as well as eliminate vital information, which is a critical reason for using fuzzy logic. Natural language can be used to record the data as well as handle the uncertainty. Any system that uses people as its sensors and source of information should use natural language, when appropriate, to identify each person's input or information. These issues and ideas led me to conclude that the use of fuzzy logic in the intelligence domain is more suitable than other probability-based approach.

In addition, fuzzy logic also excels with its ability to model complex problems through approximate behavior, its improved cognitive modeling of expert systems, and its ability to model systems involving multiple experts [Cox99]. The reduced model complexity and improved handling for analysis and uncertainty makes fuzzy logic a desirable method for handling uncertainty in a DSS. Through increasing research efforts and an extensive Japanese following, fuzzy logic systems are becoming increasingly easier to set up, use, and get accurate results. Tran and Zomorodi expect that neural networks may someday enable a computer to learn how to define the problem, set up rules, and perform any necessary fine-tuning itself [TZ94]. Other benefits of fuzzy logic based systems include reduced mean time to failure, improved mean time to repair, and easier and increased extensibility of the system [Cox99].

2.3 Information Visualization

To reiterate, the goal of this thesis was to identify an approach to enhancing the information analysts and decision-makers worked with in a DSS. This could be made possible by, oddly enough, including more uncertainty than was already present. This would only be practical if the uncertainty were identified and visualized so the users would be aware of its presence. I suggested in Section 2.2.6 that uncertainty is appropriately identified and handled via fuzzy logic. This section about visualizing information provides details that I considered when

developing my technique for expressing the uncertainty associated with the data present in the DSS.

Information visualization uses computer graphics to apply human perceptual processes to organizing and understanding data about physical phenomena as well as semantic domains to amplify cognition [CMS99] thereby providing insight rather than simple pictures. Information visualization is different from data visualization in that it involves displaying other elements such as landscapes and patterns in the data [Ack99], which typically involves thousands of data points. A common information visualization problem is determining how to use advancing graphics technology to lower the cost of finding information and accessing it once found. An issue that prevents designers from taking full advantage of advancing technologies is the human that uses the system. Approximately 9% of the population (8% male, 1% female) are somewhat color blind [Lev97]. Furthermore, those that are not color blind are still limited by our visual capacity to distinguish color, hue, and intensity differences. As the number of colors in a visualization are increased we tend to notice the smoothing and improvements, but we are generally unable to distinguish between most differences. For instance, most people do not notice the difference between their computer display using 16 and 32 bit color, which is the difference between approximately 64,000 and 4,200,000 colors.

Information visualization is a powerful approach to understanding data. Card suggests that visualizations amplify cognition in six ways: [CMS99] (1) by increasing the memory and processing resources available to the users; (2) by reducing the search for information; (3) by using visual representations to enhance the detection of patterns; (4) by enabling perceptual inference operations; (5) by using perceptual attention mechanisms for monitoring; and (6) by encoding information in a manipulable medium.

Tufte edifies various design strategies to clearly illuminate the information and message that any graphic attempts to depict. The fundamental concept for any graphics and visualization is to reveal data; through Tufte's direction graphical displays should: [Tuf83]

- Show the data
- Induce the viewer to think about the substance rather that about methodology, graphic design, the technology, or something else
- Avoid distorting what the data has to "say"
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several layers of detail, from a broad overview to the fine details
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with the statistical and verbal descriptions of a data set
- Avoid chart junk, clutter, and artistic additions
- Avoid moiré effects (visual vibrations caused by lines)

I use these design strategies and benefits of information visualization to validate the method that will be presented in Chapter 3 and symbolic augmentation it uses. I reexamine these points in Section 3.4.2 with respect to the characteristics of the visualization technique I used to express uncertainty. Section 3.4.2 explains the visualization considerations as well as how I achieved my goal of expressing the uncertainty associated with an object in the DSS.

2.3.1 Persuasive Technology

"The theory of deception is connected to the human minds perceptual and cognitive biases. Because these biases are difficult to overcome in the presence of contradictory evidence, deception can be a powerful C2W tool" [Dah96].

Fogg presents a collection of papers about persuasive technologies and our need to avoid intentionally and inadvertently using persuasive techniques [Fog99]. Common marketing techniques focus on persuasion and convincing us that we need something or we should be doing something. However, some misguided marketers apply less obvious and sometimes illegitimate means of gaining our business. For instance, the use of subliminal messages stimulating movie-goers in the 1950's, and the use of implied sex in many of today's advertisements.

The paper provides a detailed look into the intentional and unintentional side effects computers can have on our decision-making and reminds us to be responsible when designing systems involving human computer interaction. As system and program designers, we must be aware of the potential impact a tool could have, the persuasive potential of computers as well as the acceptability and vulnerability of some people. In many cases, people unduly extend credibility to their computers and the interacting systems (e.g. websites) opening them up to manipulations and persuasion.

Avoiding persuasive implications became an additional consideration of my approach for including uncertainty in a DSS. I had to ensure that the techniques I chose would not intentionally or accidentally influence a user.

2.3.2 Systems Reviewed

As part of the exploration phase of the research, I took a cursory look into several systems and computer programs in use by industry and the DOD. The first reason was to look at some of the symbology implemented and second, to determine the level of uncertainty included in the systems, if at all. The idea behind investigating the use and selection of symbols was to identify whether any graphical system had integrated or made it possible to implement the visualization of uncertainty.

The systems examined were not specialized decision analysis or reasoning applications, they were "common" systems in use throughout the military or in evaluation for implementation. I chose not to include specialized decision analysis tools because although they do express and visualize uncertainty, they are not yet standard tools for many environments. On the other hand, the decision analysis tools I had seen required extensive design and data, but were reusable for similar situations once they were built.

I looked at the Global Command and Control System (GCCS), the GCCS Common Operational Picture (GCCS-COP), the Dynamic Information Operations Decision Environment (DIODE), and the Joint Operations Visualization Environment (JOVE). I also researched evolving platforms such as the Joint Operational Planning and Execution System (JOPES), the Command Post of the Future (CPOF), the Global Combat Support System (GCSS), and the Theater Ballistic Missile Defense (TBMD). Various periodicals and research papers also provided some details about several other systems.

I found that none of the systems I examined or researched had a way for intentionally representing or capturing uncertainty in an obvious manner. In particular, I found and saw no evidence of uncertainty visualization in either of the systems. However, I was informed towards the end of my work that several databases used by DIODE do in fact record some form of uncertainty, but it is not used! The following sections provide more details about DIODE as well as some insight to GCCS-COP and JOVE.

2.3.2.1 DIODE

The Dynamic Information Operations Decision Environment (DIODE) is the "culmination of the National Air Intelligence Center's (NAIC) Information Operations (IO) analytical process" [DIO99]. DIODE is an information system that integrates intelligence about national leadership and military C2 processes, telecommunications, computer networks, and air

defense C2 networks, systems, and signals. It is designed for the intelligence analyst and decision-maker in support of the warfighter; they use it to "project hypothetical scenarios and excursions based on observed data" [DIO99]. It presents graphical and textual information relating to leadership, air defense, command and control (C2) processes as well as the fixed telecommunications infrastructure of a country [DIO99]. Although designed to support the warfighter, it can feasibly "support the acquisition and policy-making communities" [DIO99].

DIODE uses two visualization tools, Generic Logic's GLG and the US Government's OILSTOCK, to plot information compiled for a particular scenario or collection of data. OILSTOCK is a powerful mapping tool for the intelligence analyst providing several advanced tools and real-time interface. Unfortunately, OILSTOCK's power also constrains it to Unix platforms (i.e., Solaris, AIX, and DEC ALPHA). On the other hand, GLG is Java based, customizable, and designed for integration into hypertext browsers. GLG does however have several problems, some of which are discussed in Section 2.4 GLG. In addition to the charting provided by both OILSTOCK and GLG, each provides access to data using a hypertext browser.

DIODE is not a tool for the novice; its complex relationships and scenarios require direction and preparation. A DIODE user is not going to stumble on a presentation that triggers an epiphany. Furthermore, its symbology is very crude and does not meet the MIL-STD 2525 objectives discussed in 2.3.3.1. Squares, circles, diamonds, and stars seem to be its complement of icons; however, these rudimentary symbols represent a compromise between the explicit notation of MIL-STD 2525 and the variety of symbols its numerous customers use. The symbology can always be extended to integrate the MIL-STD 2525 requirements.

During the exploration phase of this research, I met with NAIC members to discuss DIODE, issues related to their field, and uncertainty [Bob99]. The meetings were used to help me understand some of the details about DIODE and some critical issues involving their work as analysts. Information included some details of the data and collection processes, corroborating

evidence, data uncertainty, as well as other reasons for uncertainty. In addition, they identified two desires they would like to have in an IDST: (1) the ability to include their own uncertainty or "risk flags" with information and objects, and (2) the ability to see the flags or uncertainty on or near the specific object. It was apparent that the analysts were aware of the presence of uncertainty and that they wanted a way to include more as well as to express it. The solution I develop to visualize uncertainty includes both suggestions and is discussed in Chapter 3.

In a later discussion with NAIC, I learned that in all but one of the NAIC databases a certainty value is recorded [Hom00]. Apparently, as new NAIC systems were evolving they were designed requiring the analyst to specify a certainty value about the data. The values 1-5 represented the validity or perceived validity of the data or information. In their implementation, the value 1 is best and essentially reflects perfect or 100% validity or certainty; the value 5 is the worst and represents the highest level of uncertainty. Although degrees of certainty are being recorded with the data, there is no utilization of the information [Hom00].

2.3.2.2 GCCS

The Global Command and Control System (GCCS) is the midterm solution and the bridge to the concepts outlined in the Command, Control, Communications, Computers, and Intelligence for the Warrior (C4IFTW[2]) concept. GCCS is an automated information system designed to support "situational awareness and deliberate and crisis planning with the use of an integrated set of analytic tools and data transfer capabilities" [DG00]. GCCS is supposed to be the single C4I system to support the warfighter from the foxhole to the command post.

The Common Operational Picture (GCCS-COP) is one of the many interfaces integrated into GCCS. The GCCS-COP or COP is a concept in which multiple applications interact to

[2] C4IFTW concept describes a fused, real-time, true picture of the battlespace and the ability to order, respond, and coordinate vertically and horizontally to the degree necessary to prosecute the mission in that battlespace [DII98].

produce a Common Operational Picture for the warfighter. This means a fused, joint and combined view of the battlespace and the ability to order, respond, and coordinate activities to prepare, support, and sustain the missions in that battlespace; essentially the C4IFTW concept. The COP capability displays land, sea, and air tracks on a near-real-time basis, overlaid onto a chart of the battlespace.

Background research in regards to the COP did not reveal any implementation of uncertainty in the visualization environment. Its contribution to eliminating uncertainty in the decision-making process is its provision of interwoven information and almost real-time operational views of the battlespace.

2.3.2.3 JOVE

The Joint Operations Visualization Environment (JOVE) is an "out-of-the-screen," stereoscopic, 3D-battlespace visual display for improved battlespace situational awareness. JOVE can provide a moving picture of battlespace platforms, units, and events over 3D terrain for a battlespace commander and staff. Visualization via JOVE allows the commander to virtually witness and better understand air, land, sea, and undersea battlespace events providing real-time situational awareness [Ack98]. JOVE allows the users to interact with "volumes of information pulled from multiple sources" [Kor97]. Through JOVE's visualization, commanders should "understand battlespace events, optimize the use of resources, and reduce the time to observe, orient, decide, and act" [Kor97].

The US Army's TRADOC Pamphlet 525-70 discusses the concept of battlefield visualization to help manage and reduce uncertainty, which they also extend to include operations other than war [DA95]. The concept and intent behind JOVE (and presented in TRADOC 525-70) is that through enhanced awareness provided by a timely and accurate view of the

environment, commanders can reduce their uncertainty by reducing the unknown thereby making better decisions [DA95].

2.3.3 Symbology

This section covers the topic of symbology to point out the constraints under which I developed the visualization technique for expressing uncertainty in a DSS. It is composed of a brief examination of the relevant MIL-STD 2525 guidelines and a recap of the symbols used in DIODE. Symbols are the fundamental elements of communication and interaction. We think in pictures so what better way to enhance communication in a DSS than with symbols.

Here are a few definitions. Symbology is the study, use, or interpretation of symbols or symbolism. Symbolism is the practice of representing things with symbols or of attributing symbolic meanings or significance to objects, events, or relationships [AHD98].

We use symbols in everything, from the pyramid on the dollar bill to the recycling symbol on consumer products to the spinning egg timer on personal computers. Symbols are pervasive in the military since they frequently capture more meaning than there is time or space to elaborate. In addition, we often augment symbols with more data to express supplementary information. The DOD provides the MIL-STD 2525 as guidance for symbology in the warrior domain.

2.3.3.1 MIL-STD 2525

Under the Defense Information Infrastructure Common Operating Environment Initiative, military computer systems, including DSS, use or are required to use the Common Warfighting Symbology provided in MIL-STD 2525. The symbol set was "designed using human factors engineering research to eliminate conflicts within the symbol sets" [WSSP99] while providing C4I symbols. In addition to symbols, MIL-STD 2525 provides coding schemes for automation and information transfer, an information hierarchy, and technical details to support

systems [WSSP99]. The design team synthesized its symbology from land-based, nautical, and aeronautical warfighting domains. MIL-STD 2525 is the primary reference the DOD uses to standardize warfighting symbology.

MIL-STD 2525 provides extensive and explicit details about the symbols, colors, icons, text, positioning, and so forth. Figure 18 provides a brief idea of the tactical symbol make-up as well as the various fields surrounding the icons used to present text modifiers. Each symbol can be composed of a frame, color fill, icon, and modifiers. Figure 18 includes a symbolic sample with modifiers (on the left), and a second image that indicates the positions (darker in color) where various modifiers can be added to the symbol (on the right).

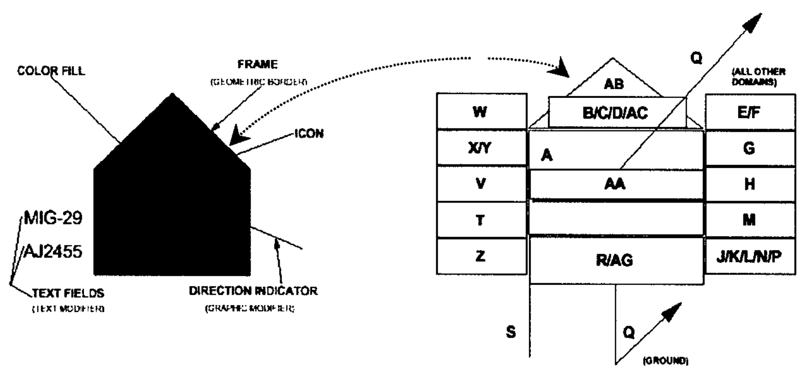


Figure 18. MIL-STD 2525 Symbol Components and Text Field Positions

The frames around the icon serve as the foundation for adding other components like the text modifiers to the symbol. The frames can also indicate meaning, when solid the physical position is certain versus dotted indicating that the position is suspected or planned. A symbol's color fill represents the affiliation of the object (i.e., friend, foe, neutral, or unknown). When it is not used, the icon and frame color are made to reflect the object's affiliation. Two types of modifiers, text and graphics, can augment various portions of the frame and icon. Some modifiers are simple graphical additions such as a directional indicator or echelon indicator while other textual modifiers might include combat effectiveness or staff comments.

Although designed using human factors engineering, I found several details that make the symbols objectionable or difficult to use. First, the textual modifiers are clear and distinguishable when enlarged like Figure 18, but they are lost in actual use when the symbols get smaller in size as shown in Figure 19. In addition, augmentative text is difficult to read among common map items; I suspect that performance decreases when users are forced to decipher the interwoven and precisely placed text. Second, the symbols themselves become indistinguishable from a zoomed-out view. There is an apparent trade-off between some legibility with larger symbols and increased separation with smaller symbols.

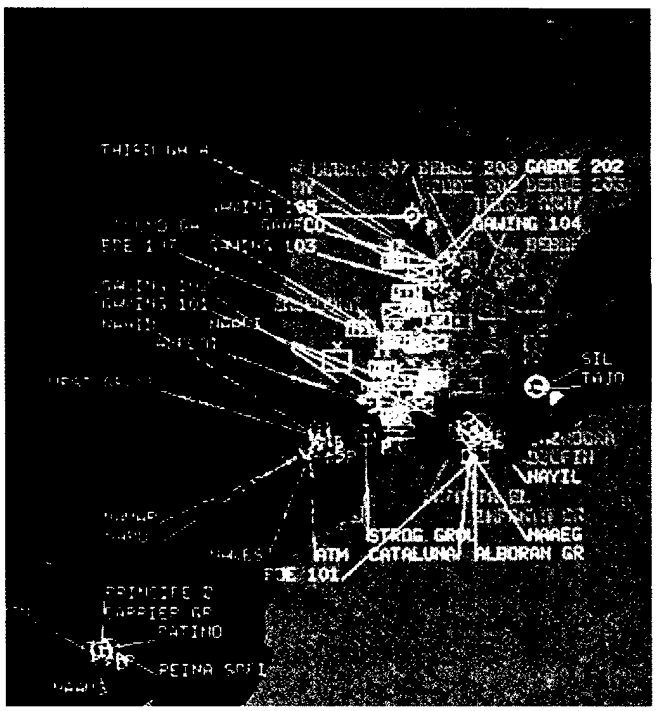


Figure 19. Challenging Symbol Use

Recently the Naval Health Research Center (NHRC) replicated and extended previous findings in a study of visual search performance. Their study of tactical displays and three variations of symbols (outline, gray-solid, and colored-solid) found that colored-solid symbols had the fastest search and recognition rates with the least confusion [Ord99]. The outlined

symbols required the greatest visual dwell time and resulted in the greatest occurrences of misidentification.

In observance of their findings, adding additional information to any icon must be accomplished in a manner that is conducive to use rather than confusing. As part of this thesis, I enhance the icon and extend the symbolism it embodies. An implication of the NHRC findings suggests the use of solid symbols or those appearing solid because they stand out against the background, particularly those contrasting with the background. In my technique for visualizing uncertainty, I use three closely placed lines to help the details stand out. In this case, I determined that the lines were less likely to be misinterpreted as a map detail as compared to one thick line that could be misread as a road. As is evident, one of my considerations for augmenting the icon and symbols will be clarity as well as conspicuousness.

2.3.3.2 DIODE Symbology

At first glance, DIODE's symbology appears very crude and lacking of intuitive inferences, it does not even meet the MIL-STD 2525 objectives. Squares, circles, diamonds, and stars appear to be its complement of items. The snapshot in Appendix D provides a few samples: blue boxes and white circles. These rudimentary actually symbols represent a compromise between the explicit notation of MIL-STD 2525 and the variety of symbols its customers use. The DIODE symbols are similar to the Naval Tactical Data System symbols, which are generally square, round, and diamond shaped. In addition, the graphical interface of DIODE is secondary to its purpose, which is to aid the intelligence analyst with analysis (and decision support) through correlation of objects, events, and processes as well as the modeling hypothetical scenarios. I used some of DIODE's symbols in my prototype in comparison against some of the MIL-STD 2525 symbols, which I established as an assumption in Chapter 1.

2.3.4 Uncertainty Visualization

Uncertainty visualization is a technique for extending information visualization by including the uncertainty about the data with the data [Cha97]. Unfortunately, most visualization research has ignored or avoided the presentation of uncertainty [AVC99]. Only within the last few years has it been actively pursued and applied as a method for handling uncertainty.

Like information visualization, uncertainty visualization pairs human perceptual capabilities with visualization to identify anomalies and intricacies that might otherwise be missed. In some systems, accuracy is as important as the data, but often lacking, in which case, visualization without uncertainty can be misleading and unreliable. I observed this in some DSS and resolved to identify an approach for including and expressing uncertainty as my thesis. I expected and found that visualizing uncertainty would aid data analysis and decision-making.

The challenge behind uncertainty visualization is the difficulty of defining and characterizing the uncertainty, hence the prior discussion of uncertainty and the establishment of the Taxonomy of Uncertainty. Other challenges include the presentation and control of uncertainty in the visualization process. Until recently, there were few methods to present uncertainty, particularly of data for large-scale 3D data sets and visualizations [AVC99, WPL96].

2.3.4.1 Uncertainty Visualization Methods

Several uncertainty visualization methods are conceptually simple, yet effective while others are more challenging and scrutinized under numerous research opportunities. These methods typically capture the expected value or event and include a representation of the uncertainty or other possible values that may occur. Most of these effects relate to one dimension of data. Although uncommon, it is possible to represent more than one dimension of uncertainty. Researchers at the University of California Santa Cruz (UCSC) developed several techniques for uncertainty visualization [WPL96], including the approach to vector fields shown in Figure 20.

Their technique reflects uncertainty in magnitude and direction using an enlarged and elongated arrow. You can get a feeling for the ocean flow, speed, and possible variances of either by simply looking at the visualization. Other methods for representing uncertainty include variance indicators, scatter plots, shading, glyphs, sliders, clouds, colors, and sweeps. Uncertainty visualization can be presented in a subdued manner to serve as a subtle reminder of the presence of uncertainty or highlighted and even exaggerated to stand out or help with analysis [Cha97].

UCSC accomplished extensive work in the field of uncertainty visualization. One of their achievements was the redesign and modernization of the taxonomy of uncertainty visualization methods. The taxonomy provides several alternative visualization methods for four different applications: radiosity, animation, interpolation, and flow [WPL96]. Unfortunately, I could not employ their examples, methods, or information in my approach to uncertainty visualization. Essentially, I had already started down a course of action when I came across their research. On the other hand, I found that some of their information supported the approach I was working on: their information suggests the use of glyphs for 1-3 dimensions of data. I realized in my exploratory phase that I was dealing with multiple dimensions of data that could each have multiple dimensions of uncertainty, which is why I was developing the Taxonomy of Uncertainty. I expected to use the taxonomy to help identify a technique for visualizing multi-dimensional uncertainty.

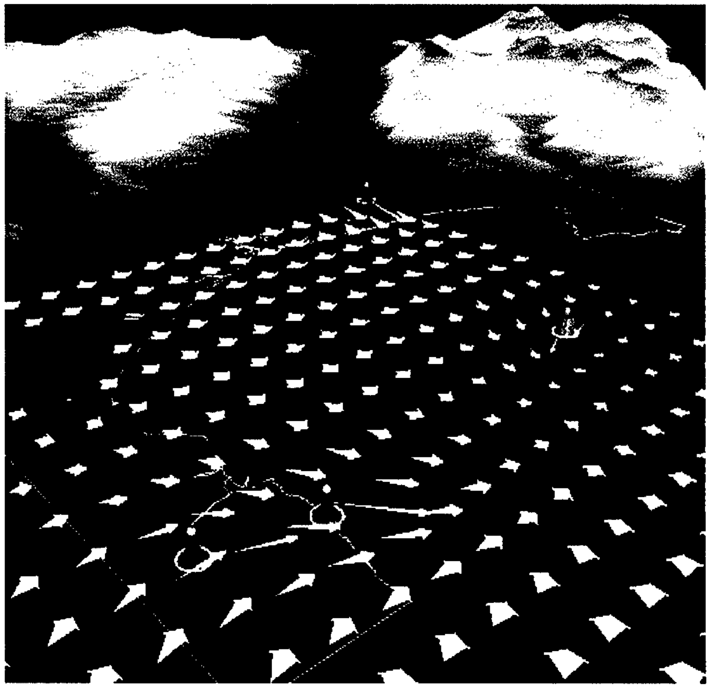


Figure 20. Uncertainty Visualization in Vector Fields [WPL96]

The remaining sections identify various uncertainty visualization techniques I considered while developing a solution to expressing multi-dimensional uncertainty in a DSS. My technique is presented in Chapter 3, specifically Section 3.4.4.

2.3.4.1.1 Tukey Boxes

Tukey boxes are one of the older methods. Tukey boxes involve drawing a box or rectangle over the interval of the occurring data values as seen in Figure 21. Lines often come out the ends of the box to points indicating the highest and lowest values that occurred in the data. The box shows the values that tend to occur while the protruding lines show the extremes. In this sample, the line in the middle of the box shows the average and the point in the box reflects a specific value such as the mode.

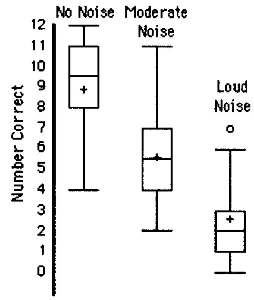


Figure 21. Tukey Box

2.3.4.1.2 Variance Indicators

An implementation similar to the Tukey boxes uses the line chart with markers indicating the beginning and ending points as well as data interval. These lack the box that shows the frequently occurring data. Figure 22 shows a line chart commonly used in the stock market, the vertical lines represent the value of the stock for one day. The vertical line indicates the range of values covered that day; each left horizontal line indicates the stock's opening price for the day, while the right horizontal line indicates the closing price. Color enhances the visualization making the red or lighter lines showing loss stand out, black and darker lines show gain.

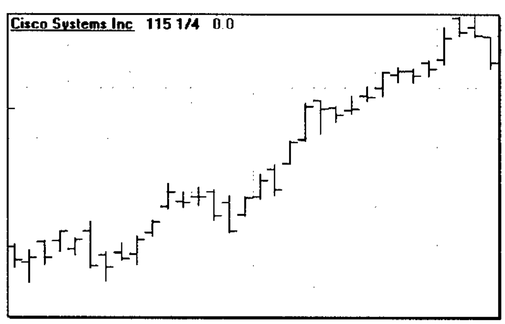


Figure 22. Line Chart

Figure 23. Candle Chart depicts another way of showing the same information as Figure 22 but in a slightly different manner. Again, red indicates loss and black shows gain. These icons are read according to color or solid versus hollow. For red or solid candles, the top of the box or candle represents the opening price while the bottom represents the closing price. The

lines that protrude from either end indicate the highs and lows during that day. The black or hollow candles reflect gain, thus the bottom line represents the starting price and the top is the closing price.

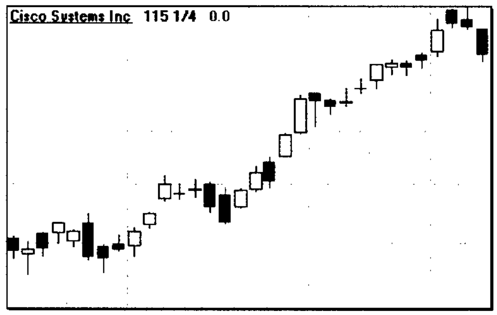


Figure 23. Candle Chart

2.3.4.1.3 Scatter Plots

Scatter plots are not necessarily designed to represent uncertainty but do so by its inherent representation of diversity. Trends can be identified in some scatter plots but explicit predictions are unlikely. Weather reports are scatter plots of millions of points commonly color coded according to a standard scheme.

2.3.4.1.4 Shading

Shading is another simple yet effective way to indicate change, variance, and uncertainty. Typical hurricane predictions will use color and shading to specify the different possible tracks and different probabilities of being hit by the hurricane. Other uses of shading occur in various areas to indicate an increased variability in range, size, or distance. Haziness is also a form of shading because the haze is accomplished using other shades.

2.3.4.1.5 Glyphs

Glyphs are symbols that represent data through visual properties such as color, shape, size, and orientation. Glyphs are also called probes, geometrical primitives, stars, boxes, and icons. They represent data points unlike icons or symbols, which refer to information, concepts,

objects, or actions within a user interface. A few glyphs and sweeps are shown in Figure 24. Glyph selection commonly, but not always reflects the data. For instance, lightning has been represented in weather reports as yellow circles, lightning symbols, as well as plus and minus signs. UCSC uses a glyph (see Figure 20), similar to an arrow to express direction and speed in one of their examples.

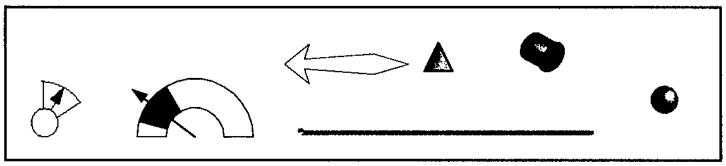


Figure 24. Sample Glyphs

Recent studies have found several ways to indicate a difference in values; hence uncertainty, by using other objects or symbols such as spheres, cones, lines, and tubes. Several varieties of glyphs allow utilization with virtually any visualization. The UCSC team provides several examples in their work, particularly in the use of arrow and sphere glyphs. Dispersed 3D spheres can indicate fluctuating fields, density, and volumes while 3D arrows (Figure 20) can indicate speed, orientation, and variances in orientation.

2.3.4.1.6 Other Uncertainty Visualization Methods

Other methods of representing uncertainty include overlaying and side-by-side comparisons, transparency, use of contours, vectors and segmenting or blurring, radial swirls and wipes. UCSC identifies several more approaches including streamlines, animations, and oscillations such that the differing views reflect uncertainty [WPL96]. Still more UCSC approaches include perturbations to represent randomized surface roughness.

2.4 GLG

GLG is an extension of Java Bean from Generic Logic Incorporated [GLG99] used in DIODE (Section 2.3.3.2 DIODE) and the design of the sample program, discussed in Chapter 3.

DIODE's use of GLG was the primary reason for my evaluation and use of this tool. I have provided a summary of its capabilities and some of its weaknesses in this section.

A significant benefit behind the GLG product is the graphical design environment *GLG Builder* that is used to create and define graphical entities, objects, attributes, and animation. These objects are used in Java applets and applications serving any purpose you desire. The GLG material presents several ideas involving meters, animation, modeling, and more to demonstrate their capabilities. A second advantage of GLG is its support for the Windows and Unix operating systems and its practical set of application programming interfaces (API), which include C/C++, Java and Java Bean, ActiveX, and a Netscape plug-in.

The GLG toolkit "replaces tedious coding to create animated graphics with an intuitive point and click graphical editor interface" [GLG99]. The GLG Builder enables you to simply draw your 2D and 3D graphical objects instead of dealing with low-level details of Java AWT and graphics. From within the toolkit, you can edit each object's attributes, define their dynamic behavior, attach dynamic constraints, and then "immediately prototype it enjoying double-buffered flicker-free animation and automatic damage repair" [GLG99].

The GLG toolkit is a component-based architecture that uses GLG drawings as its building components. As a result, all GLG graphs and other components are just drawings that are available for use whenever the GLG Java run-time engine is finished. The GLG "composite component approach" allows the editing any of the objects and parameters in a graphical editor, *GLG Builder*. Adding new functionality to an object results in a new drawing that can be used without requiring any new libraries or classes, resulting in a real component reusability. This made it possible to incorporate dynamic graphics as well as graphics based on dynamic data.

Unfortunately, GLG or its inherent capabilities behaved poorly on my computer. The GLG Java Bean or its foundation, Java, is CPU intensive and does not manage memory well. I forced explicit and frequent garbage collection, but made very little improvement to its extremely

excessive memory utilization. Running the sample program, discussed in Chapter 3, for more than a few minutes (with less than 100KB of data) quickly consumed more than 60MB of RAM. In addition, program response was very poor: simple zoom functions took several seconds and sometimes up to a minute to complete on a 400MHZ AMD K6-2 computer. While GLG earns its platform independence through Java, it may also be the reason for its very poor responsiveness. Generic Logic representatives were not aware of any problems and assured me that it does not behave as poorly on a Unix system; further testing of that nature was not a concern at this point. On the other hand, I had interesting results while attempting to integrate the prototype into a hypertext page. I found that performance and memory utilization were significantly better and negligible when using Microsoft's Internet Explorer 5.0 to access the prototype as an applet. Apparently, Microsoft's Java virtual machine does a better job of controlling and executing the prototype's source code that was founded on software designed explicitly for Sun's Java and Java Bean. Unfortunately, other issues precluded further integration into a hypertext page.

2.5 Summary

This chapter provides some insight to the issues and processes presented in the following chapters. Each section and issue I have presented is used or considered in development of my approach for enhancing information used by analysts and decision-makers by including and expressing uncertainty in DSS. Topics ranged from cognitive processes and heuristics to uncertainty then to information and uncertainty visualization. Cognitive issues and biases affect the method, symbols, and augmentation that I chose to express uncertainty. In addition, information visualization strategies and goals were used with the Taxonomy of Uncertainty to develop the technique for visualizing uncertainty in a DSS. The fuzzy logic method of handling uncertainty links directly to the approach chosen to represent the data, computation of uncertainty, and the visualization itself.

3. *Methodology and Implementation*

Many intelligence reports in war are contradictory; even more are false, and most are uncertain.

Carl von Clausewitz
On War

This chapter presents the approach I devised to including uncertainty in decision support systems (DSS) in order to provide the users with enhanced information. I organized the chapter in an incremental approach to explaining my technique for visualizing uncertainty and my approach for enhancing the information used in a DSS. There are six portions to this chapter, they start with an explanation "objects," specifically my use of the term "object" as well as how I defined the objects used in the prototype. The next section revisits the topic of identifying uncertainty and where I discuss a compromise to solely using fuzzy logic. Following that, I discuss a couple of ideas for computing multi-dimensional uncertainty associated with an object, after which I talk about visualizing uncertainty. In the visualization section, I explain the environment, visualization goals, several ideas, and then present my techniques for presenting uncertainty in a DSS. In fifth section, I layout each element to my approach for enhancing the information used by analysts and decision-makers. The final section presents and explains the prototype and its modified implementation I created to demonstrate the concept of including and visualizing uncertainty in DSS.

The methodology I propose includes a paradigm shift in more than one way. Fuzzy logic is not fully accepted in the USA let alone the military and the idea of including and visualizing uncertainty in a DSS is also radical.

I eliminated all possible references to the source code to focus attention on the concept and various aspects of the methodology rather than the program that models the concept. However, these can be made available upon request to AFIT.

3.1 Objects in General

During my investigation of uncertainty, I found it necessary to come up with a generalized term that suitably identified various items and things that were being evaluated. I chose to use the term "object" since it can be used to refer to anything tangible such as cars, tanks, switches, and agencies as well as intangible things like software, circuits, flight plans, and schedules.

3.1.1 Assumed Object Composition

In general, I assumed all objects were formed of various materials, parts, information, and even other objects in a modular sense. Since each object was identified in terms of parts and information, I also assumed it had a purpose and means to accomplish its task no matter how important or menial it was. I made these assumptions to represent and identify any object in an elementary decomposition, which can be documented in files and records delineating the intricacies of the object. DSS, like DIODE, are automated information systems that use the data provided by various databases about objects and their objectives.

I also assumed that the information about objects used in DSS includes the composition, resources, and capabilities of the object or at least some of that information. One of the features identified in my initial assumptions made in Chapter 1 was that uncertainty was recorded in the data. As such, every object is defined by a collection of information that includes uncertainty and can be examined by people and parsed by computers. Section 3.2 provides a few methods for identifying (labeling, not finding) uncertainty in a practical manner.

For the prototype program, I defined a particular arrangement for the information about each object and a method for recording the uncertainties, which is presented in Section 3.1.2. Although the organization of object data is not an issue addressed by this work, I think the method I use in the prototype could be refined and used for real data. The method I used in the

prototype is flexible, extensible, and predictable. The technique breaks the data into four areas that are designed to capture data under general categories. Within these elements I also denote a place to record uncertainty and a method to indicate whether the item or sub-elements would be used in the computation of uncertainty.

3.1.2 *Prototype Objects Defined*

Each object in the prototype is represented by the information recorded by an individual file rather than a database record. The files were Extensible Markup Language (XML) based and part of the source code I eliminated from discussion, hence minimal inclusion. Figure 25 is a diagram of the object and the classification of its information. The four mid-level elements of Figure 25 show that the information was grouped into four simple categories: object identification (object_id), key properties (key_properties), other properties (other_properties), and analysis (analysis). I chose these categories because I felt they best represented the classification of any information that I assumed might be associated with an object. As specified, the classification provides for object identification, the separation of critical and less important information as well as analytical contributions.

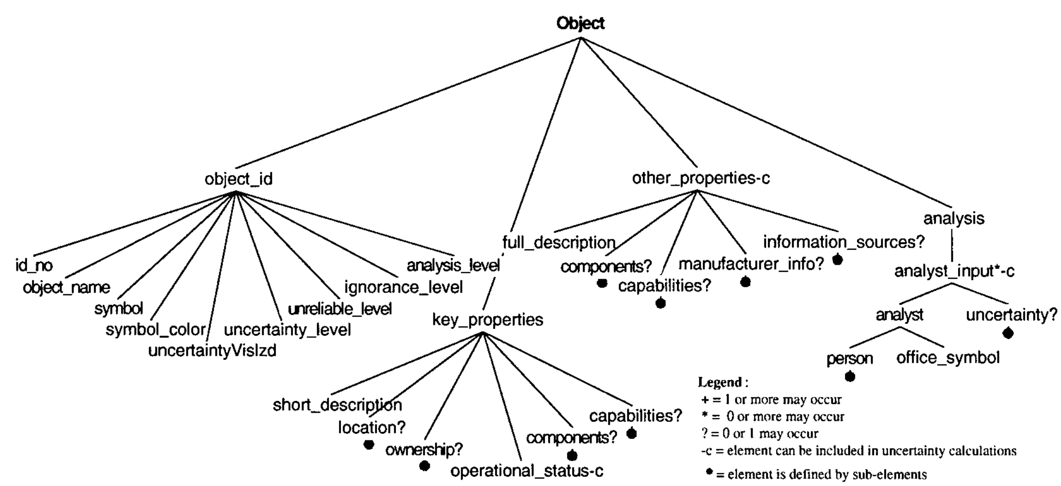


Figure 25. Object Categorization

The notation used to define the object and in the following figures is specified in the legend shown in Figure 25. Each figure in this section represents subdivisions of elements identified in Figure 25 and the data file of the object. The • (dot) below any diagramed element (e.g. person or location) acts as a virtual connector that indicates the element is defined by another sub-tree. The "-c" is notational, not part of the element name. The "-c" indicates the element has an attribute that I named "calculative," which indicates whether the element or its sub-tree is included in the calculation of the uncertainty associated with the object. Sections 3.3 and 3.5 explain the uncertainty calculation in more detail. I used a set-like form and logic based notation. The set-like form identifies the components of an element, while the notation simplifies the quantification of each component (shown in 3.1.2.2). The "?" means that there are 0 or 1 occurrences of this element or component, and the "*" means there are 0 or more of these elements. Only the leaf nodes (i.e., object_name) of a branch or tree contain data. Non-leaf nodes identify sub-trees or nodes and can contain the calculative attribute (i.e., analyst_input), but will not contain data. Data is composed of plain text, which could be words, numbers, text-based symbols, or a mixture of all. The actual data type of any element can be constrained or converted by the program using the data.

Sections 3.1.2.1 – 3.1.2.6 explain the four categories of information I used to record mock-up data for the prototype. The sections include the method I used for recording uncertainty in the data and also explains the repetition of components and capabilities, under which I record most data about an object. The more self-explanatory elements used in the prototype and shown in the diagrams, for instance short_description, location, and operational_status will not be explained.

3.1.2.1 Object ID Area

The object_id sub-tree, element, and data consist of information that uniquely identifies each object. In this case, I determined it would include an identification number, the object's

name, the reference number of the symbol that is used to represent the object, the symbol's color, and five uncertainty-related fields. The symbol and symbol color correspond to the icon and color options that are used to represent the object in the visualization environment. MIL-STD 2525 defines a plethora of symbols, their variants, and color options of which the prototype only uses 19 icons numbered 0-19. In most military battlespace visualization and planning tools, the icon color is determined by affiliation (i.e., friend, foe, neutral, and unknown). The prototype uses a few uncommon colors for diversity and as an example of the different colors that could be used for the uncertainty visualization.

The five uncertainty-related fields are related to the technique I established for expressing uncertainty, and are explained in Sections 3.1.2.6 and 3.4.4. Section 3.1.2.6 explains the uncertainty fields I used for the prototype, while Section 3.4.4 identifies how the Taxonomy of Uncertainty was used to identify a technique for expressing uncertainty.

3.1.2.2 Key Properties Area

I expected that the information pertaining to an object would be extensive and could be separated into two collections of information based upon their overall importance, such as the critical or less important information. The area key_properties is used to record the information that could be deemed critically important about any particular object. Although I created the sample data and determined what was important for the prototype demonstration, someone or something would have to decide the significance of the information and where it belonged. One idea mentioned in Section 3.3 identifies the potential for dynamic organization based upon the scenario or situation being evaluated.

The key_properties are shown below and in Figure 25. It includes information about the location, ownership, operational status, components and capabilities as well as a short description. Figure 26, shown below, illuminates the composition of several elements within key_properties.

key_properties = (short_description, location?, ownership?, operational_status -c,
components* -c, capabilities* -c)

An ownership element was included as an example of information I thought might be relevant in some situations; its composition is diagramed in Figure 26. The person sub-tree is provided with some other elements since it is a sub-element of ownership. The reuse of elements like components and person are possible by the modularity of the objects I defined.

The elements called components and capabilities were intentionally identified in both key_properties and other_properties, these and the uncertainty field shown as part of location are explained in Section 3.1.2.5, after the other_properties and analysis sub-sections are explained.

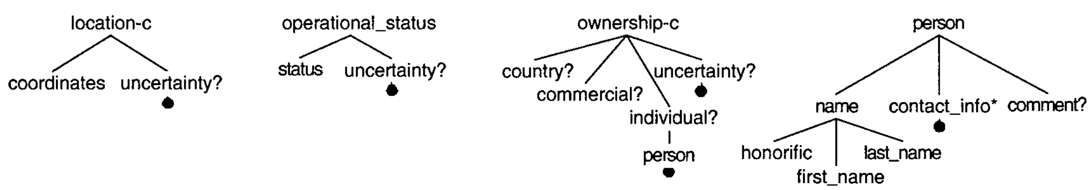


Figure 26. Visual Breakdown of Several Elements

3.1.2.3 Other Properties Area

The other_properties consists of the object's less important properties and information. By my design, this information is less critical and includes more components and capabilities as well as other information. For instance, the address of a battle tank manufacturer is rather insignificant compared to some of its other information such as its operating range. However, data that is generally less important could be an important issue to a facet of another problem, which is why I represent that it is retained in the first place. In a dynamic system, as the one sketched out later in Section 3.3 Techniques for Computing Uncertainty, the relevance of the information could be determined on demand.

other_properties = (full_description?, components? -c, capabilities? -c, manufacturer_info? -c,
information_sources? -c)

The second use of components and capabilities is intentional and records the features that are less important to the overall impact or success of the object. Take for example the dome light in a car, it is not essential, but still mentioned in the owners manual. These two fields are discussed further in Section 3.1.2.5.

Another item captured under other_properties is the identification of the information sources that defined or supplied information that identifies the object; see Figure 27. The space could be used to identify contacts, books, wiretaps, conversations, or whatever was necessary and could be used as reference for additional data, like a thesis citing a resource. Some DIODE databases record this information as a means of evidence. The element called manufacturer_info is also shown in Figure 27 as an example of data I thought was less important.

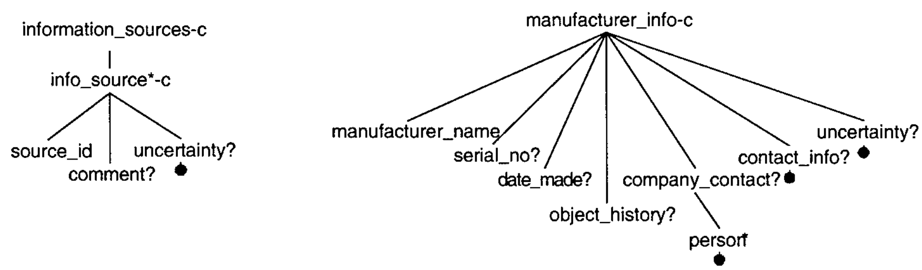


Figure 27. Information Sources and Manufacturer Information Diagrams

3.1.2.4 Analysis Data Area

The analysis category shown in Figure 25 would allow the people creating, analyzing, or working with the information to add their own input with regards to the object, its information, and sources. The reason behind this inclusion is the request identified in an interview of some NAIC analysts. I am trying to capitalize on the idea that the person most familiar with the object or its information might be able to add insight that is not apparent in the data. Additional details are provided in Section 3.4.4.1.4.

```
analysis = (analyst_input* -c)
analyst_input = (analyst, uncertainty?)
```

In this rendition, the analysis area can be composed of input from several analysts (i.e., analyst_input*). Each analyst's input can include his own uncertainty as a collection of issues (discussed in Section 3.1.2.6) about the object or its information. In this design, a decision-maker could omit all or some analytical opinions from the calculation of uncertainty by toggling the calculative that is indicated by the "–c" and discussed in Section 3.1.2.6.

3.1.2.5 Components and Capabilities

In key_properties and other_properties I use two broad categories, components and capabilities, to identify the composition of the object and its abilities. These are delineated in Figure 28 and the descriptions above it. As the plural of each implies, they can be comprised of many individual component and capability elements. In addition, each can record uncertainty about that element.

An example of this simple yet viable breakdown is evident looking at a simple classification of a main battle tank. Some of the components include but are not limited to the main gun, machine gun, engine, mud flaps, and tow cables. Both guns and the engine could be regarded crucial components that make up the tank and would be identified under the key_properties area. Tow cables and mud flaps are secondary information and could be recorded under other_properties. In addition, some capabilities might include thermal imaging, computer assisted fire control, five ton towing limit, and satellite communications. Of the four items, the towing limit is clearly one of less important attributes of a battle tank and would fall under the other_properties area. Uncertainties about any items and abilities would be recorded in the uncertainty element that could be included with each element, which is explained in Section 3.1.2.6.

```
components = (component* -c)
component  = (component_name, component_info?, uncertainty?)

capabilities = (capability* -c)
capability = (capability_name, capability_info?, uncertainty?)
```

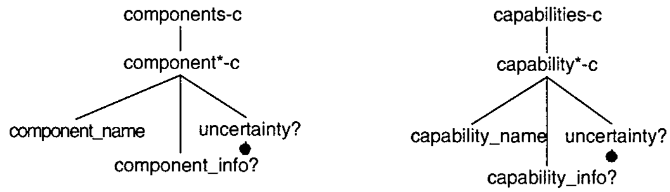


Figure 28. Components and Capabilities Diagrams

The reason I defined the components and capabilities elements in both places was because I had assumed that every object was composed of many parts and abilities, some of which are basic and generally less critical data, but still a part of the information about the object. For instance, the green paint on the side of a missile may be an insignificant component or an important capability if that paint provided stealth. As previously mentioned, someone would have to decide where the information fits. In a more dynamic system, the situation and potential use of the information would determine the relevance of information. Imagine a system that could identify relevant objects in a knowledge-store based on consideration for mission objectives. The system could dynamically search for and identify files with particular key words, then analyze those files and determine any irregularities or issues. Critical or relevant components and capabilities could be itemized under key_properties and lesser information identified under other_properties. A dynamic approach has advantages as well as added complexity in the identification of relevant objects and information.

3.1.2.6 Uncertainty Fields in the Prototype

The uncertainty element shown first in Figure 26 as part of location and ownership repeats in several other elements to capture the uncertainty that could be present with almost any part of the information. Some elements do not have an uncertainty element because I concluded that some information is generally irrelevant, in which case uncertainty about that information would not needed.

As the question mark next to uncertainty indicates, this element is optional and provides the ability to record any uncertainty associated with that particular element. The notation and diagram below specifies that each uncertainty sub-tree can be composed of many different issue elements. The issue elements were a generic way to classify the information resulting in uncertainty without explicitly labeling it. In other words, it was a way of indicating that the analyst or automated system "had an issue, problem, or concern about the information." Multiple issues provide the ability to identify multiple factors contributing to the uncertainty of any one particular element of an object. The –c indicates that this value has a calculative attribute and can be toggled for inclusion or removal in the uncertainty computation.

uncertainty = (issue* -c)
issue = (what, why, rating, modifier?, area)

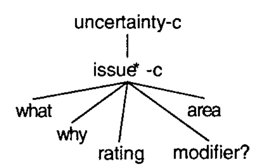


Figure 29. Uncertainty Fields in the Prototype

I intended to specify three fields within issue: what, why, and area. These fields would be used to identify and retain the information and terms indicating What was uncertain?, Why it was so? and What area of the uncertainty taxonomy it affected? Unfortunately, I was unable to implement fuzzy logic in the prototype because of time constraints, so I created two fields to record the uncertainty values: rating and modifier.

As an example of the uncertainty and issue elements, suppose you wanted to record the uncertainty about your exact position on a mountain using a commercial Global Positioning System (GPS) receiver. So the uncertainty element you would be creating or adding to would be under the location element. The issue could identify that the position is imprecise because the receiver is a commercial product. In this case, your position is what and the inaccuracy of the

commercial receiver is why. Your position may be near to the coordinates that it displays, but not exactly. For this example, the limitation of the receiver would be the cause for your uncertainty and the area field would indicate that "unreliability" was affected. We could also say that the receiver is not the problem, that the omitted signal that reduces the accuracy is the cause for uncertainty. This means that the cause for the uncertainty is ignorance by omission; therefore, area would record "ignorance." In the prototype, the rating and modifier fields record the numbers 0-100 that the uncertainty computations could use, but only if the calculative was "true."

In retrospect, I realized that a more practical approach to recording the uncertainty would have been to record the uncertainty rating with a criticality weight of the item. I could use the criticality weight as a multiplier for elements that were extremely important, they would raise the uncertainty levels during the computation of uncertainty if they had any associated uncertainty.

3.2 Specifying Uncertainty Values

As previously mentioned, I expected to include and visualize uncertainty in a DSS without needlessly increasing the complexity of the information. I realized later that part of that complexity is affected by the way the uncertainty is identified. In an earlier assumption, I specified that uncertainty was included in the data, but I did not specify how it was denoted: numerically or otherwise. The method of recording and denoting uncertainty is a significant issue to expressing uncertainty and the complexity of using those details. In this section, I provide two practical methods for identifying and quantifying uncertainty.

Our individuality as people with unique backgrounds causes challenges in the process and form of quantifying uncertainty. We interpret perceived amounts and causes for uncertainty differently, which I discussed with respect to fuzzy logic in Section 2.2.6.1. Our different interpretations and measurements indicate the necessity of identifying a semi-flexible and

imprecise method of identifying and quantifying uncertainty. The next two sections provide a brief description for using fuzzy logic and simple numbers or percentages to denote uncertainty.

3.2.1 Simple Numbers or Percentages

I dismissed earlier ideas for specifying uncertainty that used numbers because they were too similar to probability and implied a certain level of accuracy. It would be erroneous to imply precision for uncertainty, which is inherently imprecise. However, after attempting to implement fuzzy logic in the prototype, I found that while fuzzy terms may be very natural and intuitive, they were difficult to implement without a predefined library of terms, fuzzy sets, and membership functions. The prototype implements a compromise between fuzzy terms and numbers to get around the implied precision of numbers and challenges of fuzzy logic.

Numbers (0–100) were used to identify the uncertainty values I made up for the examples, which made the computation of uncertainty easy. However, I used one of four simple terms (i.e., none, low, medium, and high) to express the uncertainty to the user. Precision is intentionally removed from these values to simplify the uncertainty being expressed. The sponsor already using simple values (1-5) to denote the certainty of their data.

A practical approach to specifying the uncertainty about the information of an object could use the simple numbers 1-10 (or 1-100%) as a quantity. Higher numbers mean more uncertainty about the data or association with the object. If there is complete certainty about the information, then there is no reason to specify an uncertainty value, hence the omission of zero (0). Conversely, 100% uncertain information can be omitted; on the other hand, it could provide some advantage, i.e., as a warning. This also verges on the side of including too much data.

The interval 1-10 is more appropriate than percentages and a smaller interval. It lacks the implied precision of percentages and gives more variation than an interval like 1-4. Precise values that are available by some computation could be converted to a single digit via a heuristic

without loosing much value. The uncertainty value (1-10) could be specified by the analyst's intuition, which already occurs, or through approximations made in defuzzification processes for translating fuzzy terms to numbers.

In addition to an uncertainty measure, it would be useful to rate or weight some components and capabilities to reflect their overall importance to the object. The uncertainty of particular elements should magnify the overall uncertainty whenever there is any uncertainty associated with these elements. For instance, the location of a tank targeted for destruction should be more important and weighted more than the caliber of its machine guns.

Although I suggest converting a fuzzy term to a number (1-10), the terms that led to the number should be retained as part of the reason identifying why there was cause for uncertainty. In this approach, the numbers facilitate rapid computation while the terms facilitate human communication and intuition. The compromise between the numeric and fuzzy approach takes advantage of the advanced computers and human interaction that recognize and distinguish the uncertainties before they are added to the system.

3.2.2 Fuzzy Logic in DSS

The Taxonomy of Uncertainty identifies the roots of uncertainty and provides a means for translating common language and terminology into a unified expression explaining how and why a particular element or bit of information is imperfect or incomplete. Furthermore, the natural language used to report and generally identify an object and its information is intuitive to human use and our decision-making processes. These factors inherently facilitate the use of fuzzy logic and fuzzy terms as presented in the previous chapter. I contend that fuzzy logic is also the most appropriate way to quantify and express the uncertainties in DSS, unlike numbers that are commonly gleaned from natural language. Furthermore, fuzzy terms are stable in meaning [MF93] and facilitate use that is more dynamic.

The fuzzy terms used to quantify the uncertainty would be presented to the user without conversion or changes, unlike the simple number approach that converts the values to four terms. However, before using fuzzy logic in a DSS several issues excluding program management need to be resolved.

3.2.2.1 Fuzzy Issues Requiring Resolution

First, a library of terms, modifiers, membership functions, and equivalencies must be established and standardized. Although I found some examples, I could not find a library or mass compilation of fuzzy terms and modifiers. Second, the taxonomy of uncertainty would have to be extended to encapsulate more words, including foreign translations, and causes for uncertainty. Automation of information retrieval and categorization would be improved through the extended taxonomy and development of an elaborate library. Another issue, is the formalization of object record contents as well as the way to record issues or uncertainties contained in the information. I defined and used a particular object arrangement in the prototype to suggest an arrangement and because it met my needs for the demonstration. Other issues included the challenge of maintaining data quality by specifying and controlling who can input, specify, alter, and remove data and terms. One of the final issues is the familiarization, training and implementation of the fuzzy logic and processes. An initially difficult step will be getting users comfortable with using fuzzy logic and breaking the false security associated with hard numbers.

One benefit to using natural language and fuzzy terms in a DSS is the ability to use the data and information more dynamically. It may be possible to design a system that processes the data on demand through specialized filters where only data and information relative to a specific goal or scenario are considered in the uncertainty visualization.

3.2.2.2 Usage in Prototype

Although I discuss fuzzy logic, it and the prescribed methods were not the focus of this thesis. Fuzzy logic is a natural way to identify, quantify, and represent data, uncertainty, and the visualization of uncertainty. Validation of the fuzzy approach and the resolution of several issues must be accomplished before it is used.

A compromise using fuzzy terms and numbers is the most practical technique. A combined approach could record simple numbers for computation and retain the terms or descriptions (that resulted in the uncertainty) for the human user who wants to know more about the information. The prototype demonstrates this approach.

Within the prototype, I present uncertainty in three ways: line lengths on icons (discussed in Section 3.4.4) used in the mapping interface, bar graphs presented in the hypertext browser, and the simple values expressed under the bars. When the program calculates the uncertainty (discussed in Section 3.3) identified in the object's file, the values returned are used for the lines on the icon. These values are converted to simple terms identified as none, low, medium, and high. The browser graphs use the same simple terms as the measuring stick for the bar graphs representing uncertainty associated of any term that has uncertainty recorded, see Figure 30. The uncertainty calculation process is very rudimentary and intended only as a sample of what is possible. It computes an average of the uncertainty recorded for that object.

Figure 30. Scale as Shown in the Browser

3.3 Techniques for Computing Uncertainty

Chapter 2 revealed some insight to the many causes for uncertainty and provided the Taxonomy of Uncertainty with a detailed decomposition of the many reasons and sources of uncertainty. The taxonomy and its references indicate how and why something has uncertainty

associated with it and where various reasons fit into the causes leading to uncertainty. In conjunction with the previous discussions about uncertainty, this section provides ideas for two other questions about uncertainty to clarify other references to the use of uncertainty. Although these were not the focus of the research, they are considered part of the problem. Assuming that someone or something has already specified the data, the uncertainties, and that the object has enough information to be entered into the system, then:

- How do you quantify the uncertainty of objects with multiple dimensions?
- What does it mean to be high in uncertainty?

3.3.1 How do you quantify the uncertainty of objects with multiple dimensions?

This explanation is a rough idea and not intended to be a design approach. First, consider a simple combination of uncertainty using percentages. Refer back to the previous GPS example (end of Section 3.1.2.6), suppose you wanted to know the overall uncertainty about your exact position. A common method of specifying position is with two geographic coordinates and an elevation. Say you were 90% sure of your coordinates (longitude and latitude) and only 70% sure of your elevation. Using a simple sum and ratio you would be (90+90+70)/(100+100+100) or ~83% sure of your location, thus ~17% uncertain. A fuzzy logic approach would result in a similar finding; namely, that you were pretty sure of your location just not totally sure.

Now, consider the calculation of uncertainty for an object that involves many different pieces of information. It becomes more difficult and conceptually challenging plus it does not seem appropriate to combine the uncertainty from different types of data. Uncertainty is usually determined or accumulated between similar elements or those with some relationship to one another; however, a system that uses objects composed of many different features and information will need to merge the uncertainties in an over-arching manner. This means the

uncertainty computation will include all of the data used to identify the object, which I think can be done in a fairly representative manner.

Suppose a scenario or event called for a map that presented suspected air defenses for a country. The map would include known air defense installations as well as those with some uncertainty associated with them. In addition, suppose the uncertainty identified in some of the data was related to the firepower, communications, personnel, and detection equipment of the location. There is no explicit relationship between either of these data; to which some people would say there is no way to compute an overall uncertainty for the location mainly due to the diversity of information. I propose that there actually is a way to compute the uncertainty associated with any object based upon its inclusive uncertainties. I view this as an accumulation of the uncertainties that are compared to the worst possible cases of uncertainty that all of the information (counted or calculated) could have regardless of relationship.

The system needs a way to determine, evaluate, or identify the worst-case values that the uncertainties will be compared against. For the most part, I expect that the information about the object and the uncertainties identified would have the necessary details to identify the scales that the uncertainties are based upon, discussed briefly in Section 3.1.2.6. Otherwise, the system could include a method for looking-up the corresponding information. The look-up could be accomplished through a central data store or by including that information with the uncertainty that is being identified.

As for the computation of the uncertainty, there are three approaches to accomplishing this, the first is quick and simple, while the second is explained in some detail and the third is generalized. The first method assumes the numeric approach to denoting the uncertainty values was used. In this case, computing the uncertainty could be calculated as an average while taking into account any weights that were specified. This is the simplest approach when the numeric values use the same scale, for instance the suggested method for using simple numbers where the

uncertainty is specified by a value between 1 and 10. The worst case for these values is obviously 10. As such, the number of "marks" (i.e., calculative=true) would divide the total value that is calculated from all of the "marked" uncertainties, then the simplified scale of four values could be matched against the interval 1–10, e.g. *none* = $x < 2$ and *low* = $2 \leq x \geq 5$.

The same approximation could be applied to the second approach that uses natural language. The cumulative uncertainty could be determined by identifying, defuzzifying, and combining all of the worst case intervals for every element identified for calculation in an object's file. Then all of the recorded uncertainties would be defuzzified and combined; however, in the defuzzification the interval of uncertainty would be used, not just a single value. The processes would produce an accumulation of worst case intervals and a sum of uncertainty intervals. Dividing the uncertainty by the worst case results in a ratio that identifies the overall uncertainty.

For instance, let the graphs presented in Figure 31 represent two uncertainties identified in the data about one of the air defense sites. Suppose the graph on the left represents the reported accuracy of the air defense site over different distances. Based on the graph, the site is most accurate between 500 and 800m. In addition, let the circle represent the uncertainty about the location of a blind spot, which was reported where the site could not hit anything for some reason (say it is in a narrow valley). Suppose that the dead zone is located within their most accurate zone. Furthermore, the membership function (or person) that maps the terms that identified the uncertainty to an interval specifies that the blind spot is about 100m wide. According the graph's closest and furthest points, the worst-case uncertainty interval covers 900m. The worst case represents a situation and interval where we have data that indicates there is a blind spot, but not where it was, which means we would be completely uncertain as to the location of that blind spot.

Continuing the example, suppose other data informs us of the fuel level of the crew vehicle and its associated uncertainty. Let the data specify that they are pretty low on fuel, from a 50 gallon tank, which translated to 20% uncertainty about the fuel level being at or near the 10 gallon level. The worst-case for the fuel tank would be 50, while the approximate uncertainty = 20%*50. In the worst case, if we knew nothing about either of these elements, we would be totally uncertain or high in uncertainty. However, our perception changes since we do have some information. As for the site's accuracy, only a 100m patch is uncertain, which means a simple ratio of 100:800 uncertainty. As for the fuel, the very low identifier also reduces the uncertainty to a smaller window. Suppose some function converts the very low quantifier to 10 gallons ± 10, then the uncertainty with the fuel is limited to 10:50. Combining the current uncertainties and possible worst case interval results in a ratio of 11/95 or a little more that 10% cumulative uncertainty. In a sense, this approach is simply finding the ratio of identified uncertainty to worst case uncertainty. It is not perfect nor meant to be; remember that uncertainty itself is imprecise.

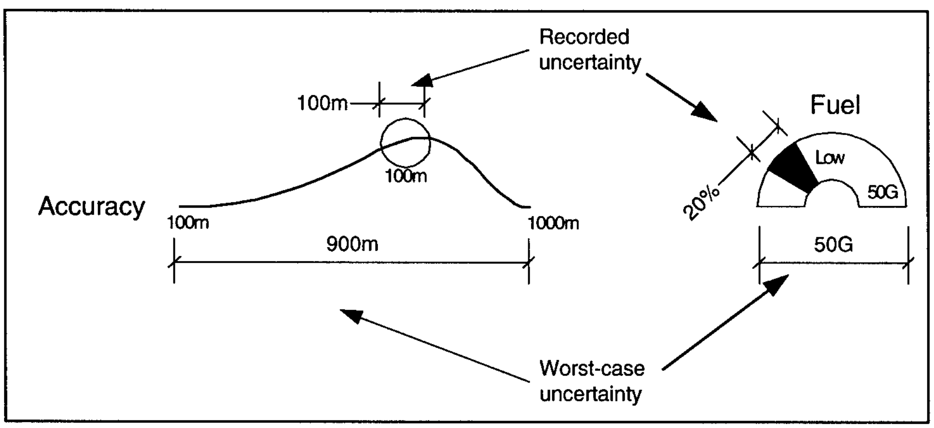


Figure 31. Some Uncertainty for Computation

The second approach uses fuzzy terms and fuzzy operations through the entire process. Modifiers of the same type can easily be combined; however, I could not find a source for combining different modifiers or different domains. For instance, combining the terms very often and seldom result in often, while high and very high result in high. The idea uses standard fuzzy

operations to combine and reduce the uncertainties down to one value. Commonly modifiers and fuzzy sets are used to create new sets [Cox99], but I am suggesting that we use the uncertainty represented by the fuzzy term as the value to accumulate. This gets difficult when trying to resolve two terms that are very different, such as "very often" and "slightly." As indicated, I could not find a source to validate my speculation, but I think it could be done.

These methods could be improved by adjusting calculations based on elements that are more significant or important than other items, as weights do in other computations. The use of weights or other values could mark or indicate these differences. Unfortunately, these ideas need development and drastically increase the complexity of the problem. On the other hand, if this were possible it would facilitate the computation of uncertainty from dynamic natural language.

A more simple and practical method for accumulating or computing an overall uncertainty uses impartial numbers, like the interval 1-10 and those in use throughout DIODE mentioned first. A cumulative uncertainty can be computed by examining all nodes or elements that have certainty values that are marked to be included in the calculation. An average would be calculated from the sum of the certainty values and divided by the number of values combined to get the sum.

3.3.1.1 Calculating Uncertainty in the Prototype

The prototype was developed to demonstrate the main concept of including and visualizing uncertainty. It uses a simple method for computing the uncertainties, which demonstrates the concept of visualizing uncertainty. The method I used calculates the uncertainty from the available information and only if it was specified to be included (i.e., calculative=true) in the computation, see Section 3.1.2.6 for more about the calculative. The elements I determined, for this model, that should have the ability to reflect uncertainty included: location,

ownership, each component, each capability, each information_source, operational_status, manufacturer_information, and each analyst.

As explained in Section 3.1.2.6, elements that could denote the presence of uncertainty had an element labeled uncertainty. The code below declares that uncertainty is composed of zero or more issues. Each issue includes fields to identify what was uncertain, why it was uncertain, the uncertainty rating (a number from 0-100), and area affected by the cause for the uncertainty (i.e., analysis, ignorance, or unreliability). The modifier field, also shown below, demonstrated the potential to express the quantities using fuzzy logic (Section 3.2.1) rather than the numbers specified above. In addition, each issue had an attribute, called calculative, that indicated if the element or sub-tree was to be included in uncertainty calculations. This element facilitates the inclusion of information without being forced to use it in the computation against the certainty of the object.

```
uncertainty = (issue*)
issue = (what, why, rating, modifier?, area)
issue calculative = (true | false)
```

The notion behind the calculative attribute also provides for the ability to allow some user interaction with the calculation of uncertainty. Consider an object and file with analytical input and the user that wants to examine the object without the analyst's opinions. Section 3.4 identifies several other ideas and benefits.

Calculating uncertainty in the prototype was a simple matter of identifying and visiting all of the fields that could be used in the calculation of uncertainty then calculating the average from the values that were present. As the various diagrams (i.e., Figure 26, Figure 27, and Figure 28) indicate, only some trees contained uncertainty. The program examined the specific trees of the three main areas of the object: key_properties, other_properties, and analysis. The entire uncertainty computation for any area simply calculates the average of the elements of sub-trees with a true calculative (i.e., calculative=true).

Only the issues of elements where the calculative is positive are counted and included thereby omitting any number of other elements whose calculative is negative. The method reflects the idea that there are enough causes with high and low uncertainties to result in values that end up somewhere in between. In practice, I expect that most elements would contain some but not a lot of uncertainty, a few on the other hand will be high in uncertainty but not enough for an averaging routine to be practical.

Although this method is very simple, its purpose was to demonstrate the feasibility of the concept. A practical solution would have to address several issues discussed throughout this chapter before implementation.

3.3.1.2 What does it mean to be high in uncertainty?

I could not find only one answer to how or why an object would be "high in uncertainty." The process or idea of applying a general quantifier to anything that is decomposable into multiple relationships and elements may not always be appropriate, but people do it every day. Consider the idea of approximating your uncertainty about being delayed enroute to work. Most people in most cities do not worry about these issues because there is little or not cause for a delay and little or no uncertainty. On the other hand, as a driver in Los Angeles, California, I was often uncertain about the delays that were frequent and often unexplained.

A trade-off between accuracy, generalizing, and acceptability is always present, yet resolved relatively easily by people. Referring back to the traffic uncertainty, I was generalizing and simplifying many of the events that could occur to cause a delay and losing a certain amount of accuracy by not finding statistical trends. I did not need a trend to tell me that I had a good chance of facing several delay-causing factors and I could roughly approximate the areas where problems would occur. So as I traveled to work, I had a very rough idea that I was highly uncertainty about making it to work on time, but this was an acceptable estimate. Acceptable,

mainly because I did not need a precise result, I needed a general idea. My point is that although there might be ways to determine very specific amounts of uncertainty, it is not necessary most of the time. Having a simple approximation of the uncertainty, whether high or low, can be enough for some people to make a simple decision.

Four possible explanations for an object to "be high in uncertainty" come to mind.

- Having one or more critical (measured) elements with a lot of associated uncertainty
- Having many elements with a significant amount of imprecision or lack of data (uncertainty)
- Having a generous mix of elements with uncertainty, for any reason ranging, between medium and high
- Having most or all elements with some degree of uncertainty

The general lack of a precise quantity of uncertainty demonstrates that a specific formulation is difficult to identify. There are many ways to interpret collections of things including uncertainty. Neither explanation for high uncertainty seems better or worse than any other; actually, each is a representation of a lot of uncertainty. The different opportunities for interpretation are not an impediment to the concept of expressing high uncertainty, but different ways to end up with similar results as in different ways to accumulate value.

High in uncertainty has many interpretations and requires an association with an object and event to further illuminate the meaning. On the other hand, I believe it is possible to generalize uncertainty to express a certain message of doubt that can be associated with an object. This thesis proposes the presentation of a symbol that simply indicates the presence of uncertainty regardless of the situation, object, and its contents. Although an approximation of the uncertainty is displayed, the object's file and data must be examined by the user to determine the specific issues and uncertainties, digging down into the data.

3.3.1.3 High Uncertainty in the Prototype

Quantifying uncertainty in the prototype was simplified to a level that merely demonstrates the concept of visualizing uncertainty. As discussed in a previous section, the method computes the average from the elements with a positive calculative, and uses the average as the uncertainty. With this process, high overall uncertainty results from the computation of elements, in the data, with medium to high uncertainty. Therefore, in the prototype, an overall high uncertainty indicates that the data reflects many elements with individually high uncertainty for one reason or another.

The resulting value of the uncertainty is not computed from the data alone. In the prototype, the analytical input also contributes to the overall uncertainty; thus, a high level of uncertainty expressed through analytical issues could also be the cause for the increased uncertainty. By way of the visualization technique I established and used in the prototype, an icon that indicates high uncertainty reflects the values specified for inclusion in the uncertainty calculations and analyst opinions, which could be disabled through the calculative. The only way to find out why there was high uncertainty is to look at the data itself, preferably through a method that aids the identification of the uncertainty. The prototype employed hypertext and bar graphs in conjunction with contrasting colors to help the information and uncertainty stand out.

3.4 Visualizing the Uncertainty

The method for expressing uncertainty is another substantial issue to the approach for enhancing the information in the DSS. A complementary consideration of expressing the uncertainty was to reduce the complexity of including this imperfect information. The emerging approach provides a way to help someone deal with uncertainty in his work, and is accomplished by expressing and visualizing the measurable parts of uncertainty. Visualizing the uncertainty

helps the user see it as a characteristic of his job, which can be extended by providing the user with the ability to add to and manipulate that representation.

The use of natural language and fuzzy logic is one ingredient to helping the user realize the presence of uncertainty in their job; the next step is to depict that uncertainty. However, in conjunction with raising the awareness of the uncertainty that was already present in the system, my goal was to improve the analyst and decision-maker's performance. Improving their performance meant improving their perspective and required a way to include more relevant, albeit somewhat uncertain, information in the DSS, which is facilitated by my approach and uncertainty visualization that is discussed in this section.

3.4.1 Visualization Environment

The visualization environments this work concentrates on are those decision support systems used by analysts and decision-makers within the military. Presumably the environment uses computer based mapping and Geographical Information System (GIS) tools as a platform for decision-making. Therefore, the tools include maps and details commonly associated with maps such as indicators of communication and transportation lines as well as natural and man-made features. Systems of this nature can include planning, logistical, and battlespace visualization tools, such as DIODE, JOVE, and JOPES. MIL-STD 2525 identifies some symbology used in this approach and the demonstration program. Section 3.4.3 provides additional details related to the symbols used in the prototype.

The concepts, methods, and issues of this approach and the visualization of uncertainty were focal points of my research and validation efforts rather than the software. As such, the documentation excludes most details related to source code and software. The prototype I designed represents the graphical mapping and data presentation portions of DIODE (see Section

2.3.2.1). The evaluation of the prototype involved user interaction with both portions to glean validation and comments from the intelligence analysts' perspective.

3.4.2 Visualization Goals

Suggestions and strategies for information visualization by Card and Tufte in Section 2.3 include several characteristics of good visualization techniques. I refined those characteristics to support this particular problem space and identified visualization goals that would reflect an appropriate presentation of uncertainty. Section 3.4.4.3 explains how I met the following self-imposed goals for the visualization of uncertainty:

- Keep it simple making it as intuitive as possible, presenting clear, legible symbols that visualize measurable parts of uncertainty.
- Make it non-intrusive by revealing the data through layers with the least user interaction and through integration with data terminology and descriptions.
- Provide for user interaction by rendering the information dynamically and providing the ability to toggle the visualization of uncertainty.

In addition, MIL-STD 2525 established several preconditions in-line with several standards of the DII-COE. The MIL-STD requires that future DOD visualization environments use the symbols defined by the standard, which also specifies that symbol colors reflect affiliation (i.e., friendly, foe, unknown, and neutral).

3.4.2.1 Visualization Possibilities

The next step to visualizing the uncertainties, associated with an object, was to identify a way to include and express it in a manner that was intuitive and stood out among other map items. Map items were a concern because the assumed DSS platform incorporates a mapping tool that uses GIS icons and symbols to represent the objects in the environment.

Unfortunately, most uncertainty visualization techniques involved and dealt with only one or two specific dimensions of data and uncertainty. On the other hand, the Taxonomy of Uncertainty was under development and the complexity of uncertainty was apparent: I was already aware that I needed to include or accommodate multiple dimensions of uncertainty. Six reasonable alternatives for enhancing an icon to extend its symbolism were (1) adding text near or around the icon; (2) adjusting the translucence of the object; (3) providing an auditory response to a mouse over event; (4) adding information, lights, or icons to the edge of the window; (5) adding glyphs, symbols, or graphs near the icon; and (6) adding glyphs, symbols, or marks to the icon. These ideas and their faults are explained in the following sections.

3.4.2.2 Textual Considerations

MIL-STD 2525 provides textual modifiers for its symbols. Section 2.3.3.1 describes the textual utilization and the possibilities for confusion. Textual modifiers meet few visualization goals and increase the visual complexity. After examining some examples like Figure 19, I determined it would be difficult for users to identify uncertainty easily if it expressed in the tiny, precisely placed text next to the icon. Text was not a viable solution.

If uncertainty must be represented in a text based system, then one could consider changing the font, intensity, boldness, slant (left or right), and other textual options.

3.4.2.3 Adjustable Translucence

Object translucence was not a viable option either. Although icon translucence could be adjusted to alter the visibility of an object, it has higher potential for being misinterpreted. Conceptually, if we made the icons less visible as the uncertainty increased, then objects that are less visible (more translucent) than others could be missed. It also seems possible to mistakenly correlate translucence to the object's existence or location. Translucence could also erroneously reflect the idea that "what you don't see won't hurt you," which is what is happening in DSS

today: details are being ignored or omitted because they are imprecise. However, this does not mean the uncertainty does not exist. The failure to include uncertainty, objects and information with uncertainty, is something I am trying to overcome rather than support. This option was undesirable because it provides several opportunities for misinterpretation and seems capable of expressing only a few dimensions of uncertainty.

3.4.2.4 Auditory Responses

I also eliminated auditory responses as a solution. In general, sound is not ideal for an environment where multiple users could interact with various objects and icons on the same system. With auditory responses, every person collaborating over one view would have to follow the pointer at all times or miss an auditory cue to a visual event. Furthermore, sound requires additional hardware and senses, which presents the opportunity for interference by external disturbances. Auditory stimulus like this should be constrained to environments with one-to-one interaction between human and device.

3.4.2.5 Window Dressing

Several common tools present extraneous information on the frame or window of the application. Although it is possible to present more information in a clear and free area, it is not always the best way for all information. Consider, for example, the common hypertext browser or word processor; both add information to the bottom of the window such as the page number or address. In either case, the user breaks focus and looks away from their work to see the data. Clearly, the interruption is not always drastic; however, the DSS environment is one where the "big-picture," focus and concentration is required. Forcing a DSS user to break their concentration to look at the edge of the window for some basic information, like a generalization of uncertainty, is undesirable. On the other hand, this is a viable method for layering information access by presenting specific details about an object such as position or source files.

3.4.2.6 Adding Symbols Near the Icon

It is possible to add symbols or marks near the icon as an alternative to placing the information on or near the window edge. Unfortunately, I found that adding symbols and graphs near the icon and others actually clutters and diminishes the view more than it improves it.

3.4.2.7 Directly Augmenting the Icon

The best idea was direct augmentation of the icon. Augmentation occurs by adding symbols, glyphs, or marks directly to the exterior frame. Symbols or glyphs would be added to the outer area of the icon frame to abide by the MIL-STD 2525 specification of not altering the symbol inside the frames. Using augmentation, the alteration would have to imply or infer an approximation of uncertainty without requiring the user to process or read additional information.

The most viable alternatives for adding information directly to the icon included clouds, sweeps, rays/lines/fuzz, and shading, some of which are shown in Figure 32. Using some simple drawings, I was able to determine that sweeps and vectors were not viable solutions because they could not address the complexity of uncertainty while maintaining simplicity and legibility. Sweeps tend toward direction or a range of values capturing at most two dimensions thereby requiring many to adequately express multiple dimensions. Adding clouds or shadows naturally elicited concern, by simply showing the icon with a ring of fog or shading implied something was different or strange. However, these too were inadequate; shading looks like a background or simple depth effect or misread as an enlarged object. Clouds or fog lack precision and when shaped to indicate zones they still seemed out of place. The simplest visualization, and most legible used the lines or fuzz painted directly to the exterior of the icon and radiating away from the center, like Figure 33.

Figure 32. An Icon and Some Uncertainty Visualization Ideas

Multiple rays, lines, and fuzz could individually express various types or causes for uncertainty; however, they could also lead to confusion and lack individual line legibility as shown in Figure 33. Although the numerous lines can indicate an abundance of data and uncertainty, they also become distracting and the lines are individually indecipherable, particularly when smaller. For clarity and legibility, the rays or fuzz need to be organized into a useful format, this is discussed further in Section 3.4.4 Organizing Gnomon.

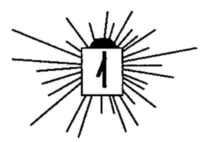


Figure 33. An Icon with Fuzz

On the other hand, the plethora of rays could be used to express individually specified dimensions of data where the individuality is not as important as the overall change to the visualization. Consider the possible impact of the visualization caused by adding data to the object. The image in Figure 19 clearly indicates increasing clutter, and as such, we can expect that adding more to the symbols will continue to increase the clutter. Fluctuations in the overall fuzziness of the icon would stand out as knowledge changed. This visualization technique could be useful for an environment of a few icons with playback or data filtering abilities. Appendix C identifies the 19 icons used in the demonstration program.

3.4.3 Gnomon Fuzz

The terms fuzz and lines seemed to lack merit and needed renaming; I named them gnomon, pronounced "NO-mun." From this point on the fuzz, rays and lines, shown in Figure 33

or Figure 35 (like legs on a spider), will be referred to as gnomon and gnomon fuzz. Gnomon is any kind of pointer that indicates a value by casting a shadow. In this case, the shadow is fuzz and the gnomon points out an anomaly, which is the presence of uncertainty as well as an approximation of that uncertainty.

3.4.4 Organizing Gnomon

I selected gnomon as the most viable symbol to express multi-dimensional uncertainty, but it needed to be organized so that its use in a visualization environment would be functional. This is where the Taxonomy of Uncertainty came into being and its importance increases. The taxonomy that identified the roots of uncertainty also fostered an approach to visualizing multi-dimensional uncertainty. The following section explains how I organized the gnomon into a more useful display using the taxonomy.

Uncertainty must be paired with an idea, concept or another word to be effective, e.g. investment uncertainty or uncertain prognosis; they mean virtually nothing independently. Therefore, the taxonomy is ultimately functional in terms of the uncertainty associated with something, whether it is an idea, object, time, person, or event. In other words, uncertainty correlates to something that is uncertain; uncertainty does not exist without being linked to something tangible or not. "I have uncertainty," says nothing and is useless. This means that the uncertainty visualization of anything will involve two points: the icon and the uncertainty. Uncertainty visualization, see Figure 34, will identify "What's uncertain?" and "Why or how is it uncertain?" Clearly, "What" identifies the object, information, or parts of them that have uncertainty associated with them. "Why/how" identifies the causes or reasons for that uncertainty.

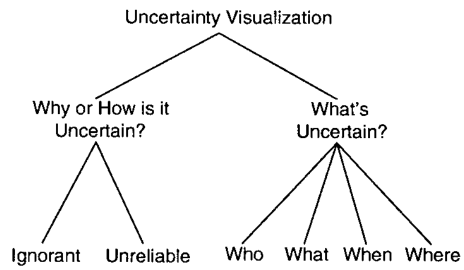


Figure 34. Uncertainty Visualization Breakdown

The statement "Tomorrow's weather is uncertain" uses uncertainty too loosely and provides only a spat of information. Conversely, "The chance of rain, tomorrow, is uncertain due to the unusually dry weather" is more precise and informative. The second statement clearly indicates "what's" uncertain as well as "why" it is uncertain. The items or elements that could be uncertain or may have uncertainty associated with them and make up "what's uncertain" are summarized in four simple words: who, what, when, and where (see Figure 34). These four inquiring words capture virtually all objects, things, information, data, and knowledge, which are represented by the icons in the DSS. The "how or why something is uncertain" provides details and information that explain the reasons and causes for the uncertainty, fundamentally composed of ignorance and unreliability (Figure 34), which are expressed by the gnomon. These points provide a natural way to express uncertainty associated with an object, by organizing the gnomon fuzz in a manner that reflects the significant details of the taxonomy.

3.4.4.1 Icon Quadrant Breakdown

After comparing several simple icons, described in Appendix B, I chose to use a simplified but distinct implementation of the gnomon fuzz. I found that the icon could be divided into quadrants, each augmented by three same-length lines radiating from the center (gnomon). This organization of quadrants and gnomon was the most legible: each quadrant and the gnomon appeared distinctly identifiable. By using quadrants with gnomon in each, the visualization can express four collections of information. Most importantly, the uncertainty is displayed in two

ways: first, by the presence of any gnomon, which indicates an issue and some uncertainty exists. Second, by using one of the quadrants I can express an approximation of the total or overall uncertainty that is associated with the object.

If one quadrant reflects the cumulative uncertainty, the other three could reflect the underlying reasons for the uncertainty: ignorance, unreliability, and analytical input. Table 4 defines the arrangement of quadrants and Figure 35 depicts a sample icon with added notes.

Table 4. Quadrant Decomposition

Quadrant	Title	Description
Upper right	Uncertainty	Represents the overall uncertainty associated with the object and identified in remaining quadrants
Lower right	Unreliability	Identifies the amount of unreliability associated with the information and its sources
Lower left	Ignorance	Identifies the amount of ignorance associated with the information and its sources
Upper left	Analytical Input	Identifies additional analytical issues or uncertainty indicated by analysts or authorized persons.

Table 4 and Figure 35 designate that uncertainty and the two main causes for uncertainty (ignorance and unreliability) use three quadrants to express their values. The fourth, Analytical Input, satisfies a consideration of the DSS: including the intuitive uncertainty or concerns of the analyst. Analysts using DIODE identified a desire to indicate, in some manner, that they had some uncertainties about the information or object. Section 3.4.2 identified several characteristics for successfully presenting uncertainty in a DSS. One such goal was the ability to provide user interaction with the symbol and uncertainty visualization. I accomplish this by facilitating the inclusion of analysts' intuition, opinions, issues, and concerns as a variable in the overall uncertainty. The sub-sections of 3.4.4.1 provide additional information with respect to calculations leading to the approximations presented via gnomon.

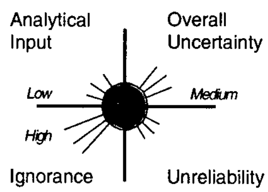


Figure 35. Icon Quadrants*

As shown in Figure 35, each quadrant of gnomon is created by three lines of equal length and width equally dispersed from the center of the icon or symbol tending more to the quadrant's center. This arrangement satisfies the concerns and requirements of the symbology and the visualization of uncertainty: clarity, simplicity, and legibility. Appendix B identifies several other ideas I considered for using fuzz.

In keeping with simplicity and fuzzy logic, four noticeably distinct lengths define the possible values that each quadrant approximates. The values are simply *none*, *low*, *medium*, and *high/long*. Each approximates the value identified in a particular quadrant; for instance the gnomon in Figure 35 depicts *low* unreliability and analytical input, *high* ignorance, and *medium* uncertainty. If we were talking in terms of probabilities, we could say none = (0, 0.2], low = [0.2, 0.5], medium = [0.5, 0.8], and high = [0.8, 1.0).

An alternatively way to use the gnomon, would to individually manipulate and specify the lengths of each of the lines in the unreliability and ignorance quadrants. The ignorance and unreliability quadrant could be designed so that there was a line for each of the areas or types identified in each: ignorance would have four and unreliability would have three. Individually the unreliability lines could represent *credibility*, *acquisitions/exposition*, and *limitations*. The ignorance quadrant might present *error*, *irrelevance*, *omission*, and *unknowable*. This method provides more insight to the causes for the uncertainty associated with an object than only using the two types. Unfortunately, it requires addressing issues of placement. As the icons get smaller

* The cross hairs are only present to support these notes.

some views of an object and its gnomon could present the situation where one or more individual lines are not shown and the user is unable to distinguish what the remaining line(s) represent. Then again, these are approximations so a little imperfection may be tolerable. This seems like an appropriate area for further examination.

3.4.4.1.1 Uncertainty Quadrant

The upper right quadrant is used to display the overall or accumulated uncertainty that is associated with the icon. I refer to it as the uncertainty or overall uncertainty quadrant. The value of the gnomon presented here is determined by combining the values of the other three quadrants (ignorance, unreliability, and analyst input).

Alternatively, we could treat the uncertainty that is associated with the object and analytical input independently. In this case, only ignorance and unreliability values will contribute to the uncertainty quadrant value. The combination in both cases is relatively simple using simple addition or the fuzzy logic "and" if fuzzy logic was being used. The substance behind the uncertainty quadrant value occurs in the other three quadrants.

I chose to include the analytical input with the ignorance and unreliability based on the idea that an analyst's opinion can be as valuable as the information identifying the object. An analyst that has been working with the details related to a particular object may have insight contrary to the contributing information.

Intuition and the mind are impossible to replicate, yet difficult to omit. When a computer system provides a summary or detailed report it is based on facts, inferences, logic and various other rationalizing processes. An analyst, on the other hand, provides the potential to gain and impart insight through synergism and perception that is fine-tuned with experience and that which no computer could mirror, yet. By facilitating analytical input, we can include that insight in the visualization and the overall uncertainty.

3.4.4.1.2 Unreliability Quadrant

This quadrant, on the lower right, reflects the uncertainty associated with an object due to unreliability. The Taxonomy of Uncertainty (see Section 2.2.5) identifies many causes and reasons explaining why an item or information can be unreliable and belongs to this area.

Depending upon implementation, the information and uncertainty might be labeled as it is added to the object's data (or file) or dynamically filtered to determine its characteristics and relevance to the issue being evaluated. Either way, combine the values of the elements identified for inclusion in the computation of uncertainty using fuzzy or basic math methods as required. Section 3.3 describes the calculation of uncertainty associated with different qualitative descriptions and domains.

After computing the uncertainty due to unreliability, a subsequent process must adjust or set the gnomon length according to the computed value. As mentioned, the length is constrained to one of four values (none, low, medium, and high) that approximates the uncertainty, which is displayed in this quadrant.

3.4.4.1.3 Ignorance Quadrant

The ignorance quadrant (lower left) is similar to the unreliability quadrant except that the calculations are relative to uncertainty that is associated with an object for reasons that cause ignorance. The taxonomy (see Section 2.2.5) identifies many causes that explain why an item or information belongs to this area.

The same process that computes the value for uncertainty due to unreliability computes the amount of uncertainty due to ignorance value that is recorded in the file. Like the unreliability process, a subsequent process adjusts the gnomon length for the ignorance quadrant according to the computed value. Its length is also constrained to one of four values (none, low, medium, and high) that approximates the uncertainty, which is displayed in this quadrant.

3.4.4.1.4 Analytical Input Quadrant

This quadrant (upper left) reflects the combined opinions and uncertainties specified by the analysts contributing to the object and its data. There may be several analysts involved in the intelligence gathering, refinement, and analysis processes surrounding an object and its information. Each analyst may have different concerns that would presumably be useful. As such, each analyst should be allowed to contribute his opinion.

However, as the number of analytical issues increase so does the uncertainty directly related to their input. This issue and many more like increasing the uncertainty weight or value of one item or issue over another must be resolved before implementing such a system. As with the previous quadrants, Section 3.3 describes the calculation of uncertainty associated with different qualitative descriptions and domains. As previously mentioned, a subsequent process must adjust or set the gnomon length according to the computed value.

3.4.4.2 Accessing the Data

One aspect of many visualization environments is the ability and process of accessing and presenting the data relative to the visualization. Several strategies for presenting and organizing the requested data; however, it is not as difficult to talk about the information that contains uncertainty. A fundamental notion behind including uncertainty in any system is making it stand out, obvious, unmistakable, and non-intrusive. Extending this idea from a visualization environment of objects and icons to a textual context is possible. One issue of some monochromatic textual environments is the inability to modify or augment the text such that it is noticeable.

Regardless of how the data becomes available to the user, the information representing or containing the uncertainty described in the previous sections must be identifiable. One method for indicating the difference between normal data and the data with uncertainty is to change various formatting attributes of the text. Using bold face to mean one thing, italics to mean

another, and underlines yet another. The font alterations can be notable when those are the only changes; however, in a document with other rich text formatting these uncertainty indicators could go unnoticed.

Another technique employs the dynamics, portability, and power of the hypertext browser. HTML, Dynamic HTML, eXtensible Markup Language (XML), other markup languages and browser plug-ins make the browser a powerful, platform independent tool for exchanging information. Applying various scripts, styles, and dynamics integrated into most browsers can change a simple textual report into a multimedia and interactive event. A simple textual document (with some formatting) can incorporate colors, sounds, flashes, highlighting, charts, lines, pictures, and more, expressing its message or projecting its uncertainties. I use this method for presenting information about an object accessed from the demonstration program.

3.4.4.2.1 Presenting the Data in the Prototype

In the prototype, each object has its own data file that contains all of the data for that object and any uncertainty associated with it. Each object and file is independent of the others, data is not shared between files nor cross-referenced. I chose to use XML for the object data files for several reasons, but have removed all examples of it from this document in order to focus on other aspects of the research. For future reference, the six reasons were: its innovative approach to recording information, its forthcoming global standardization, its potential as a truly global and platform independent form of data exchange, its increasing interest and support by industry, its relative simplicity to use, and finally, its lack of evaluation at AFIT.

Section 3.1.2 explains how each object of the demonstration program was broken down into four primitive categories: object identification (object_id), key properties (key_properties), other properties (other_properties), and analysis (analysis). XML enables this breakdown through its tagging and document object model.

The prototype used the capabilities of Internet Explorer to employ several different methods for ensuring that areas with uncertainty stood out and supported the user rather than debilitated him. As Internet Explorer loads a data file, the format of information being displayed is applied dynamically. Dynamic formatting occurs according to the instructions specified in its accompanying XML style sheet. The formatting style I specified presents the information with uncertainty in a distinctly colored table with a bar graph depicting the uncertainty that is present or associated with a particular element. Figure 36 shows an item with a contrasting box that highlights the presence of uncertainty. The emphasis is used to point out the reason for uncertainty as well as to provide an approximation of the uncertainty that is associated with the noted element or issue.

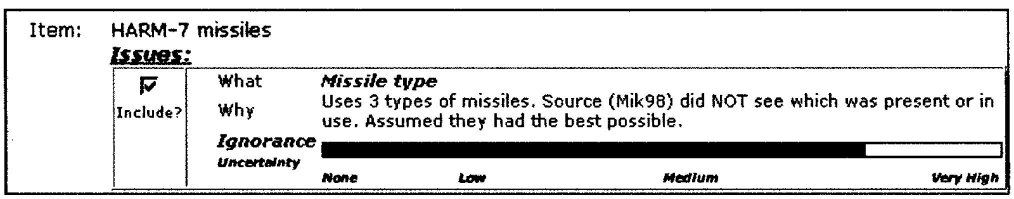


Figure 36. Uncertainty Presented in Data

With the style sheet and the dynamics of HTML, every issue of uncertainty is presented in a frame similar to Figure 36. XML, XSL, JavaScript, JScript, ActiveX, HTML, and DHTML were used to dynamically present the information and augment the presentation making it easier to use. Section 3.6 provides more information about the prototype.

3.4.4.3 Visualization Goals Accomplishments

This approach to visualizing uncertainty in a DSS accomplished the goals I stipulated in Section 3.4.2 Visualization Goals in several ways. Although presented here, these findings are validated through a prototype that was evaluated by the sponsor, and discussed in Section 4.3. The three goals are itemized then followed by an explanation of how each was satisfied.

- Keep it simple making it as intuitive as possible, presenting clear, legible symbols that visualize measurable parts of uncertainty.
    - As Figure 35 and Figure 36 show, quadrants are distinct and the gnomon fuzz stand out while remaining distinct enough to avoid blending with the background. The gnomon lengths are generally discernable and represent the uncertainty that is recorded in the data. By computation, the gnomon specifically represents the elements that were flagged for inclusion in the uncertainty computation of the demonstration program.
- Make it non-intrusive by revealing the data through layers with the least user interaction and through integration with data terminology and descriptions.
    - The prototype design, discussed in Section 3.5, appears to be intuitive and easy to use via mouse clicks. The data terminology is not explicitly integrated; however, the gnomon stands out thereby inferring the presence of uncertainty, which is displayed by the gnomon. On the other hand, the idea for providing access to the data was implemented in the prototype and provides access to the actual data defining an object. The prototype reveals a variety of the object's data in layers through three events: mouse placement over the icon, and mouse left and right clicks. The desired user interaction is discussed, but not thoroughly implemented in the prototype due to several challenges of the programming interface.
- Provide for user interaction by rendering the information dynamically and providing the ability to toggle the visualization of uncertainty.
    - The ability to toggle the visualization of uncertainty gnomon was included and provided some user interaction. Finally, the information is rendered on demand as it is used; real-time visualization would be ideal. Real-time visualization could support dynamic low-level information changes to an object.

This approach is founded on numerous considerations intended to reduce complexity and enhance the speed of comprehension while providing the user with additional information. At this phase of evaluation, I have stepped through various portions of the solution, which seems obvious, simple, and intuitive. But we need to realize, that the solution is more apparent due to the integration of the taxonomy, which makes the causes for uncertainty less nebulous. The approach even starts to appear too simple, its simplicity is noticeable when you consider what I perceive the cognitive load of this approach is to a user.

3.4.4.3.1 Cognitive Load of the Uncertainty Visualization

The approach discussed for visualizing uncertainty seems to fall into the following itemization and respects the typical cognitive limitations of the average person. The items that challenge the cognitive load's magic number seven (± two) are

1. That uncertainty features (gnomon) are available and augment current symbols,
2. The uncertainty visualization is broken into quadrants of three equal-length lines,
3. The gnomon lengths (none, low, medium, high) are relative to the object's uncertainty,
4. The upper right quadrant represents the overall uncertainty,
5. The lower right quadrant represents the unreliability of the associated information,
6. The lower left quadrant represents the ignorance of the information or source,
7. The upper left quadrant represents the analytical comments and opinions,
8. Detailed information about the object and its uncertainty is available by "digging-down" into object's data, and
9. The visualization of gnomon can be enabled or disabled for each object, each attribute, by areas, and globally.

3.5 An Approach to Enhancing Information in DSS

The major portions of my approach to including enhanced information in the DSS have been presented over the last several sections and culminate to the general diagram in Figure 37. The basic idea to enhancing the information the decision-makers and analysts must contend with is to include uncertainty in a practical manner without unduly increasing the analyst and decision-maker's burden. The previous sections presented those ideas as well as some other issues, some of which are demonstrated in the prototype discussed in Section 3.6. The approach that will enhance the information that analysts and decision-makers work with is pulled together in a final review of the ideas incorporated.

My approach, which facilitates including uncertainty in DSS, can be broken down into five main ideas. These ideas involve the processes for (1) recording the uncertainty; (2)

identifying objects and elements relevant to the situation at hand; (3) computing the uncertainty based on those elements identified; (4) visualizing the uncertainty; and (5) providing user interaction with the uncertainty.

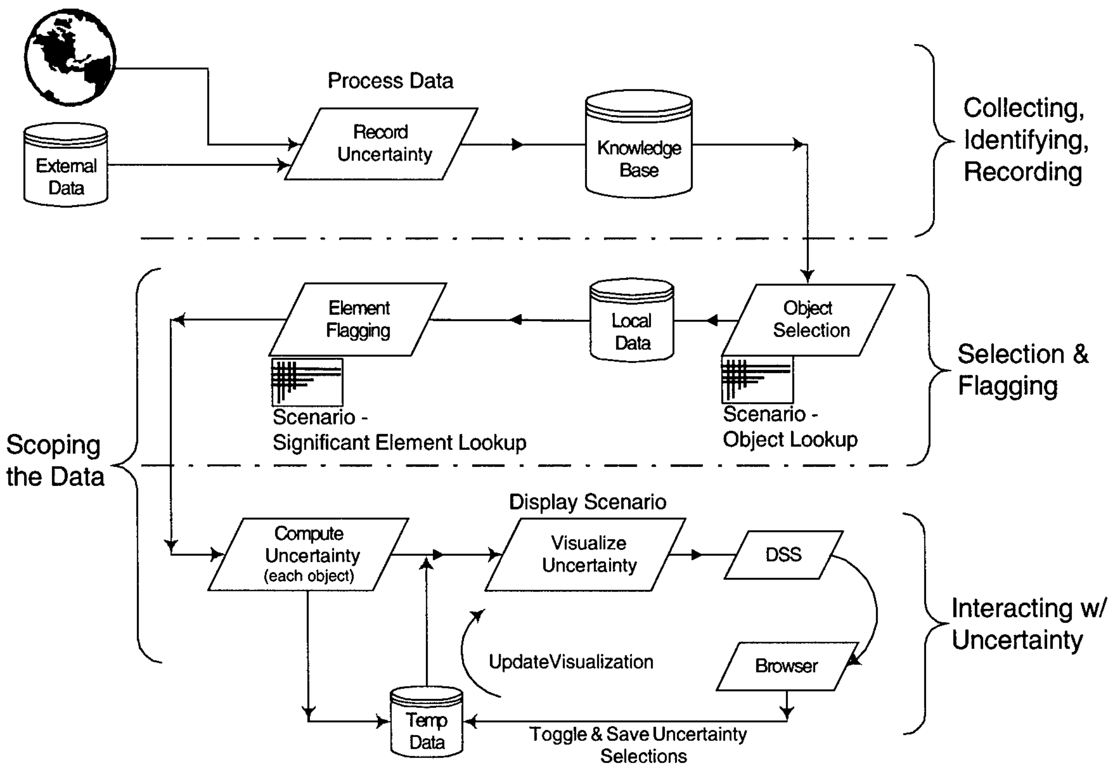


Figure 37. An Approach to Enhancing Information via Uncertainty Visualization

3.5.1 Recording the Uncertainty

A major portion of this approach is the process of identifying and recording the uncertainty that is associated with an object and its data. This was not the focus of my research; by an early declaration, I assumed indicators of uncertainty were recorded and available. However, I have provided some ideas for recording uncertainty since this task is the foundation for the entire process.

In review, the desired approach would retain the natural language used by the source to report the information and data. Although my earlier thoughts suggested the solitary use of

natural language and fuzzy logic, I determined that a compromise was necessary because of the difficulties I had with fuzzy logic. One of the issues I looked into was the availability of fuzzy intervals and membership functions for the numerous fuzzy terms that could be used to describe something. Unfortunately, I could not find a library of terms and intervals that were available for use, which led to the idea of a compromise.

The compromise involved simple integers (1–10) in conjunction with natural language. I believe the system should save the natural language capturing the uncertainty (or a summary of it) for later human use. In addition, it should use a simple number (1–10) to specify the amount of uncertainty that is encapsulated in the information defining the object.

The numbers simplify computation while the terms simplify human comprehension. Through the inclusion of numbers or other terms, we can also specify weights or ratings that are useful for distinguishing especially important components or details. Retaining the natural language and terms also facilitates future growth into more dynamic systems. Eventually, computing power and formalized fuzzy terms could provide the foundation for expressing, manipulating, compiling, and combining the data using only fuzzy logic. The process for dynamically computing the uncertainty that is associated with some information becomes increasingly complex as the number of terms grows.

3.5.2 Identify Objects and Elements Relevant to Situation

The second process in this approach involves the selection routines that would determine the objects and elements of each that should be included in the scenario. Although this process was not part of my research, it is important to note its necessity in that it effectively narrows the scope of the search space, problem domain, and the objects examined. The following steps identify how I envisioned this system might work.

First, a process determines which objects will be used in the DSS. A look-up table and user selection could specify the objects that would normally be included or evaluated as part of the scenario. Then collection processes would seek out and retrieve the desired objects from a specialized or global knowledge base storing the information locally. Obviously, the scenario or situation under evaluation will dictate which objects should be included for evaluation. The second process could use another cross-reference or pre-defined specification to indicate which particular data elements should be flagged for inclusion in the computation of uncertainty. The DSS should have a reduced collection of objects, which have been each pre-filtered and prepared for uncertainty computation and visualization when these process are finished.

3.5.3 Compute Uncertainty of Elements Identified

The next step of this process is the computation of uncertainty. As previously mentioned there are a few techniques for computing the uncertainty of similar or related terms. However, the difficult task is actually applying those formulas in an acceptable manner to uncertainty associated with dissimilar elements and data. Without a fuzzy logic approach, that I was comfortable with, I presented a compromise that used simple numbers for recording and computing the uncertainty.

The uncertainty of any object would be calculated from the information included in its file or records that were flagged for inclusion in the computation. In the prototype, I called this flag the calculative. External factors and information were not considered as part of the current method because other related issues (complexity and delivery) have not been explored. On the other hand, if the system is bountiful in its information gathering and storage there will still be a significant amount of irrelevant data associated with each object (in their records) that pre-filtering and flagging would remove from computation.

3.5.4 Visualize the Uncertainty

Visualizing the uncertainty associated with each object is the fourth process and involves the presentation of gnomon in graphical DSS and the use of bar graphs in a browser. Gnomon directly augments the icons and is separated into quadrants that express the multidimensional uncertainty. The quadrant and gnomon combination presents a visualization of an approximation of uncertainty that is clear, legible, and practical for the analyst and decision-maker.

Bar graphs extend the uncertainty visualization to lower levels of evaluating the data associated with a given object. This was necessary because I expected that DSS users would eventually want or need to access the data. The bar graphs are used in two ways, as a simple indicator of the approximation of uncertainty associated with that specific element and as a "signpost" that would help the user locate the elements with uncertainty.

It would be counter-productive and only a partial solution if the uncertainty was not presented in a manner that made it stand out. The last thing I wanted to do was include uncertainty then force the user to plod through torrents of data to find those uncertain elements included in the file. With that in mind, I employed different coloring schemes to further reduce the chance that the areas with uncertainty would be missed, see Figure 36.

3.5.5 Provide Interaction with Uncertainty

As previously mentioned, there could be an enormous amount of irrelevant information in a given object's file. As such, a pre-filtering process must mark the elements, of an object, that should be included in the calculation of uncertainty. Presumably, some information and details will be unnecessary or undesired for consideration and users should be allowed to adjust the selection of the elements being included in the computation of uncertainty. In other words, they should be able to set and change the flags as they see fit. This also facilitates the evaluation of elemental uncertainties and how they affect the overall uncertainty that is associated with the

object. This would allow the analyst and decision-maker to toggle the uncertainty visualization in such a way that lets them to see more and less uncertainty.

Selection can be implemented in various ways. The prototype used tables of tables, which made it possible to toggle large blocks of uncertainty as well as individual elements. This allowed the user to include or remove varying degrees of uncertainty allowing them to see how changes in their information would alter a situation. Although discussed in the next section, the prototype employs a browser interface for user interaction and an overall uncertainty graph that the user could use before updating the DSS.

Supporting further user interaction with the data and the DSS, the system should allow the user to add their own information and uncertainty to the local data. Conceptually, they analyst or decision-maker in the battlefield will have access to more current information as well as a different perspective. With this in mind, the end user could conceivably provide information to higher echelons as well as to the analysts who are maintaining the primary knowledge base.

3.5.6 Application to Other Domains

This generalized approach to enhancing information can be modularized and employed in different stages of information processing. The most likely modules would be the processes for identifying, categorizing, and recording uncertainty, followed by a process for identifying and selecting objects as required by each situation. The task of flagging specific elements within each object's records could be separate and subsequent to the object selection process or combined with it. The uncertainty computation process should also be a distinct module. On the other hand, the uncertainty visualization via gnomon requires integration with the DSS environment. However, the display of data, uncertainty bar graphs and ensuing interaction with uncertainty toggles can be separated onto itself and another process as well as another interface.

The techniques I have discussed were designed in consideration of DSS using multiple, complex objects having multi-dimensional uncertainty and is unlike common uncertainty visualizations that engage very few uncertainties and thousands of simple data points. Furthermore, the ideas and concepts for visualizing the uncertainty were based on the premise that uncertainty itself is imprecise and therefore the uncertainty visualization will also be an approximation, rather than explicit.

Finally, it is worth noting that the approach to enhancing information I presented is not constrained to a specific domain or environment, which means it can be applied to almost any domain involving uncertainty. It is well suited to objects with multi-dimensional uncertainty and DSS platforms used to aid the decision-maker.

3.6 Prototype or Model Program

To validate the approach and techniques presented over the last two chapters, I created a prototype that demonstrated the potential employment and viability of these ideas. Fundamentally, the approach facilitates the inclusion of information and objects with uncertainty thereby presenting a more detailed and informed view of a situation or scenario. The program, scenes, and information reveal these benefits by presenting a workspace similar to DIODE (discussed in Chapter 2) that facilitates the inclusion of uncertainty. A supplementary survey completed by evaluators and discussed in Chapter 4 corroborates the concept and utility of including more information by visualizing uncertainty.

3.6.1 A Likeness to DIODE

Unlike the approach that was discussed in a more abstractly, the prototype was intentionally designed resembling DIODE. The likeness to DIODE would provide a familiar backdrop and circumstances to an innovative concept for the sponsor and analysts who would

evaluate the program. The analysts, that were going to evaluate the prototype, used DIODE regularly. By using a likeness of their system, I eliminated unnecessary details related to understanding the environment, display, and scenario allowing the evaluator to focus on the concept of uncertainty visualization being demonstrated. Unlike DIODE, which provides additional features (discussed in Chapter 2), the model program presents static information following and the evaluators followed a simple script (described in Chapter 4) that demonstrated the different aspects and benefits of including uncertainty. Other non-sponsor evaluators, who would be unfamiliar with DIODE, were familiarized with the idea behind the graphical DSS. The non-sponsor evaluators were expected to be composed of decision-makers or students of decision analysis. I anticipated that non-DIODE users would still contribute through their evaluation of the uncertainty visualization techniques as well as providing a fresh perspective to decision support environment. They would be able to focus on the uncertainty visualization as if it was being presented on some unspecified platform for visualizing the battlespace. The decision-makers would provide opinions directly related to the complexity of using a system embodying uncertainty and the difficulties of making sound decisions on systems with and without enhanced information.

The program specifically resembles the mapping portion of DIODE, in which icons and connections are displayed on a map. The map and icons represent the environment (country and objects) involved or related to the selected scenario. DIODE employs two applications to display and operate the GIS-map interface (OILSTOCK or GLG) and for accessing the data (Netscape's Navigator). I chose to imitate the GLG mapping portion rather than OILSTOCK. GLG was under evaluation by NAIC as an alternative to OILSTOCK and its Java foundation made it possible to work on a Windows personal computer rather than a Unix-based workstation.

For the most part, DIODE's mapping interface is duplicated, but its technique for accessing the text-file-based data that is associated with a specific object was completely revised.

It takes several steps, in DIODE, to access the data, which is a simple, colorless, textual listing of the information. Their form for selecting and toggling the objects that are displayed on the map is better, but not likely to be dynamic. The prototype also provides access to the data; however, requiring less effort and using Internet Explorer (instead of Netscape). Internet Explorer supported XML while Netscape did not; the prototype's mock-up data files were XML. Section 3.4.4.2.1 provides a brief explanation of the technique I employed for accessing the data. As the reference to the mock-up data implies, I did not develop nor incorporate any means of demonstrating the first two processes identified in Section 3.5. The omitted processes were (1) recording the uncertainty and (2) identifying objects and elements relevant to the situation at hand. On the other hand, the mock data identified and recorded uncertainty as suggested. In addition, the object selection was simulated by the configuration files, which referenced two files that identified the objects (filenames as shown in Table 5) for inclusion and the connections between them. Internal to each file and by way of the calculative attribute, I flagged the elements, in each object's file, that would be part of the uncertainty computation.

3.6.2 Mapping Tool and Gnomon

Java and GLG were used to implement the GIS-mapping tool that displays an almost featureless, birds-eye view of the battlespace, see Figure 38. Generic Logic Inc provided a temporary license for their enterprise edition *GLG Builder* used to create and define graphical objects and more, see Section 2.4. I modified a GLG network traffic demonstration creating a revised interface and palette in addition to the Java application source code, which I used to enable the mapping interface. I added several capabilities in addition to the 19 icons that represented different objects. The list of objects and some sample icons are provided in Appendix C.

In the prototype, the country in the birds-eye view is the US and only includes country and state boundaries in addition to the icons and links representing the objects and connections between them. The prototype is started by invoking a batch file that references two files: G#_facilities and G#_links*, which identify the facilities and links between them to be displayed. The program loads the map and any objects, in the order identified, in the referenced files. The facilities file is a list of files, shown in Table 5 on the left, while the connections between them are listed in the links file, shown on the right of the same table. Table 6 denotes the format of each file. The method for indicating and loading the files (and objects) by the prototype is similar to DIODE, but an improvement over its data file specification. When DIODE tells the GLG map what to load, it includes the coordinates, symbol number, and color. Admittedly, it seems like DIODE is using the extra information externally because GLG is not a completed implementation; it is under evaluation. In the prototype, these details are unnecessary and omitted because they were made part of the object's data, which is evaluated by the GLG application. This simple change reduces the exchange of information between the selection processes, data source, and the mapping tool. For instance, when loading the facilities, the prototype only requires the list of object filenames and the path to each. Table 5 shows the sample contents to one file listing facility objects and one of interconnecting links.

Table 5. Sample GLG Map Reference Files

Facilities File	**Links File**
\GnomonData\Radar_SavanahGA.xml	1 3 0 2
\GnomonData\SAM_TallahaseeFL.xml	2 1 1 5
\GnomonData\SnglRckt_BruinswickGA.xml	2 3 5 5
\GnomonData\OregonCntr.xml	0 1 9 15

Table 6. Reference File Formats

Facilities File	**Links File**
Path\data_file E.G. Path: \GnomonData\ data_file: OregonCntr.xml	1st Site, 2nd Site, Link Type, Link Color 1 5 3 2

* The # is a placeholder for a number 0-4; e.g. G3_links.

The first two columns, of numbers, in the links file correspond to the facilities, which are numbered, as they are being loaded. The numbering starts from 0 with the first object loaded. The program loads the specific facility files in the top-down order listed. The program parses and extracts any required data from each of the specified files; the data includes the object name, location, icon symbol, color, and values for the four quadrants of gnomon (if present). The information gleaned from a file is used to define and characterize a specific icon in the mapping tool; the icon also retains the path and filename in a hidden attribute. Section 3.1.2 explains the file organization and content.

If uncertainty values for the gnomon lengths were not present in the data, the contents of the file are examined and the uncertainty computed. Uncertainty values might be blank because the file was not preprocessed or values were not retained from a previous use. Sections 3.2.2.2 and 3.3 explain the uncertainty calculations. Once the uncertainty of the object and gnomon length is calculated (even if 0 for things with no uncertainty) the icon is added to an array of objects that will be shown on the map. The program processes the links after loading all of the objects. Like the objects, the connections between icons are added to an array that saves the links and their interconnecting details. When the processing of objects and links ends, the map is updated with its information and the icons and connections are displayed.

Although most of this document avoids programmatic details, the following two sections (3.6.2.1and 3.6.2.2) provide a few minutiae about the operations occurring behind the scenes and within the program.

3.6.2.1 Using the Map

The map responds to left and right mouse clicks. Left clicks tell the program to display a blue dialog box in the upper right corner; the new window (see Figure 38) provides some information about the point or item selected. The dialog box can reveal the object id, facility

name, location, and a file path if these details are present and attached to the object. Right clicking on anything other than an icon is the same as a left click. Right clicks on an icon causes the program to invoke Internet Explorer. As the prototype calls Internet Explorer, it provides Explorer with the path and filename of the object, which Explorer automatically processes as input. Sections 3.6.3 and 3.4.4.2.1 provide more information regarding the data and its presentation by way of Internet Explorer.

Key features of the demonstration involved presenting uncertainty and the ability to toggle its display globally as well as individually. Other standard features included the ability to toggle icon visibility on/off, toggle link visibility on/off, toggle map visibility on/off, toggle label visibility on/off, and the ability to load four other demonstrations without reloading the program. Figure 38 shows a snapshot of the prototype with various features pointed out; Appendix D provides a better view. The figures in Appendix D include objects that further demonstrate the use of gnomon to indicate the presence of uncertainty. The white circle in Figure 38 is one of the symbols used in DIODE, which has no uncertainty, while the other one in the northwest is taken from MIL-STD 2525 and augmented with gnomon fuzz. The snapshots in Appendix D include several plain blue boxes and white circles of DIODE and several gnomon-bearing icons.

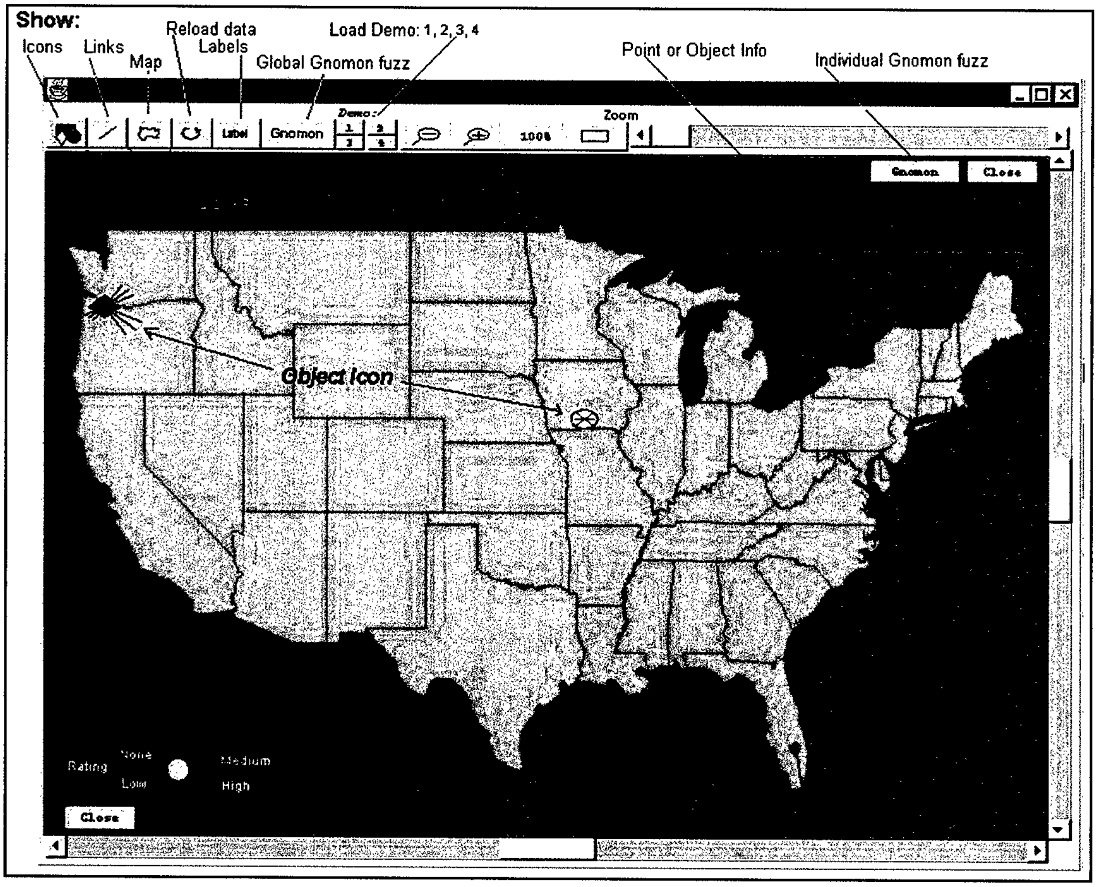


Figure 38. Prototype Snapshot

The prototype exploits several techniques to provide information to the user. With a mouse-click anywhere on the map, a dialog box displays some basic data about the object or point selected. When an object is the focus of the mouse-click, the data includes the object's name, location, and file name. In addition, a button is provided to toggle the gnomon visibility of the individual icon on/off, Figure 38 includes a dialog box in the upper right corner that shows this button. Otherwise, it displays the coordinates for the point and state name, if applicable. The interface also displays object names as the mouse floats over their area. Finally, if you right-click on an icon the application invokes Internet Explorer and uses the path of the object's data file as input for the browser. Eventually, Explorer loads the files and presents a closer, more detailed look at the information related to the object as explained in Section 3.6.3.

3.6.2.2 Behind the Scenes of the Mapping Tool

At different stages, objects and icons are manipulated to reflect the toggling of a feature or data being loaded. Accessing the objects and icons defined using the *GLG Builder* was rather easy considering the complexity involved. The explicit and clear modularity provides for simple data references.

One time saving feature, was the ability to group icons together, then by constraining various attributes to one another I could cause global changes by making a single change to one of the grouped icons. For instance, the visibility attribute of all icons were constrained to one another. To toggle the visibility of the icons simply meant the program only had to change the visibility value (shown below) of one icon, unlike some systems that would have to update the value for all displayed elements.

```
/Icons/Icon1/Group/IconVisibility
```

Likewise, toggling the gnomon globally meant changing its `GnomonGroup/factor` value to 0. In GLG, the factor is a multiple by which the gnomon fuzz is magnified, it is normally set to 1. Toggling the gnomon individually required a change to the visibility of the gnomon, not the icon.

Changes to the gnomon length occurred by changing the `scale` for the appropriate quadrant. Each quadrant for the first icon is demonstrated below.

```
/Icons/Icon0/Group/GnomonGroup/UncrtntyGnomon/scale
/Icons/Icon0/Group/GnomonGroup/UnrelblGnomon/scale
/Icons/Icon0/Group/GnomonGroup/IgnrncGnomon/scale
/Icons/Icon0/Group/GnomonGroup/AnalysisGnomon/scale
```

Unfortunately, the ability to change the gnomon line width was not available. According to Generic Logic, their early implementation of Java Bean did not include the ability to change the line width from within the application. Worse yet, the application reset any line widths specified in the toolkit during design to the default width value 1, which made the lines and gnomon very thin. Slightly, thicker lines could make the gnomon more noticeable. In addition to the standard attributes, GLG allows you to add and define custom variables. This allowed me to

include and store the file name and path of the object that Internet Explorer used to access the data after the details were loaded into the map. This was the best way to link each object with its data.

3.6.3 Displaying the Data Via the Browser

The prototype uses Internet Explorer (a.k.a. the browser) as its interface for displaying the contents of the object's file and interacting with the uncertainty. I used a style sheet and JavaScript to format, style, and enhance the object's data that was loaded into the browser from an XML file. The object's information is displayed in a manner designed to make the issues or uncertainties **stand out**, see any of the next four figures. Figure 39 demonstrates another technique used to visualize the uncertainty present in the data, it shows the graph of the accumulated uncertainties for each quadrant. In retrospect, I realized that by displaying or using the same icon and gnomon here, which represented this object on the map, I could reduce possible confusion caused by transitioning from an icon with gnomon to cumulative bar graphs. Figure 39 and Figure 41 also show the contrast used to make the uncertainties and approximated values represented by the graphs stand out.

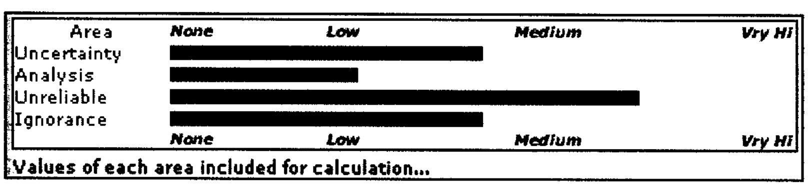


Figure 39. Another Method of Presenting the Uncertainty

Figure 40 shows the four main sections of information that define every object used in this demonstration, discussed in Section 3.1.2. At the top of the screen, the object's identifying details are clear, as is the graph presenting the cumulative view of the uncertainties that are associated with this object, like Figure 39. Using DHTML styling, I can hide the information contained in several areas, in Figure 40 a blue title bar highlights these areas. Although not a

major feat, it provides a nice way for the user that wants to tuck away information. Post-it notes mark areas that can be or have been concealed; Figure 40 shows them concealed and Figure 42 exposes two areas. Moving the mouse over certain places exposes tips and information or details that explain different parts of the display. Shown in the upper right of Figure 40, in a yellow box, is one such mouse over event, which indicates the value of the top bar, Uncertainty; it says, "Rated @ 75%." Each one of the little features used in browser improves the presentation of the information and the user's interaction with it. Appendix E includes several more images that cover the information displayed for one object.

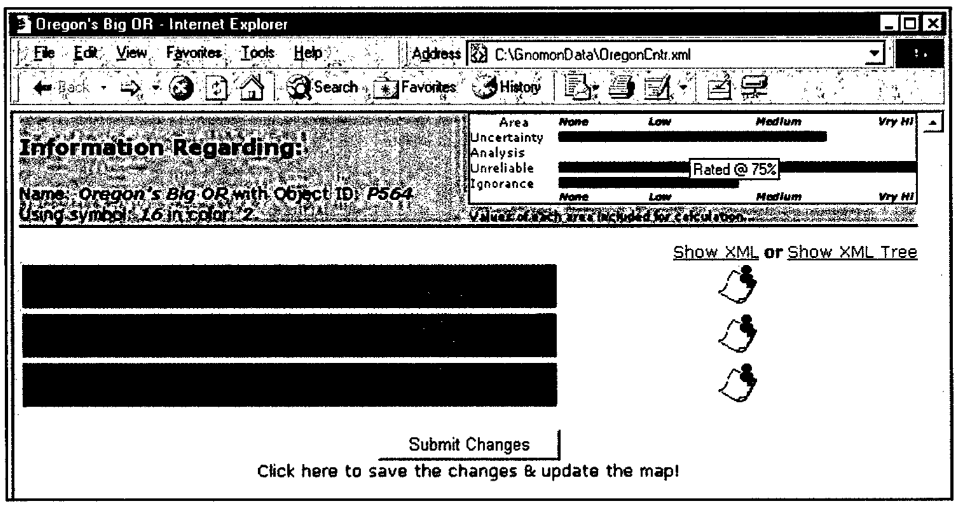


Figure 40. Browser View, Areas Hidden

One of the more notable features is the top portion of the window that shows the identification and overall uncertainty; it is designed as another layer that always stays on top. This layer provides the user with a readily available view of the uncertainty. The always-on-top layer allows the user to move to any point in the file and still have a visual depiction of the overall uncertainties for each quadrant and cause for uncertainty.

To interact with the uncertainty selections, the user would toggle check boxes, as shown in Figure 41 and Figure 42, to see how much the uncertainty of a particular element affected the overall uncertainty that is associated with the object. Unfortunately, this was not finished. The

checkmarks do however reflect the element's calculative, a checked box means the value was true and the element was used in the computation of uncertainty.

The information about the object is broken into three other areas (shown in Figure 40), each highlighted blue: *Analyst input and opinions, Key properties, and Other properties.* The data presented throughout the document is in tabular form to help organize the sections and information as well as making it possible to toggle large chunks of uncertainty. Several other features that make it more readable and useful are presented below.

Lighter colored boxes make *Issues* or uncertainties stand out among normal data, see Figure 41. The bar graphs, shown in Figure 41 and others, paint the level of uncertainty recorded for each item identified with uncertainty. The bar graph indicates the amount of ignorance, unreliability, or analytical uncertainty that is associated with this specific item. The source area (analysis/ unreliability/ ignorance) is indicated on the left side of the bar. Shorter bars indicate less uncertainty and are therefore better! For instance, Figure 41 shows a lot of ignorance is present or identified with respect to the missile type.

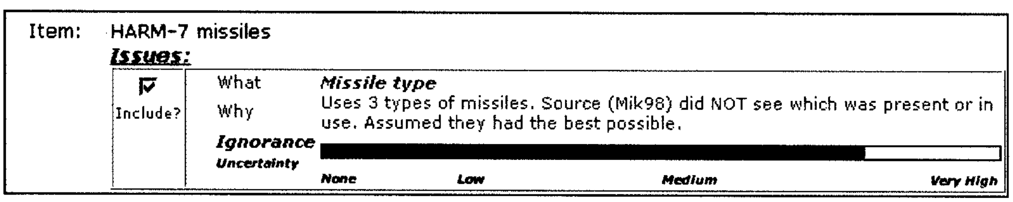


Figure 41. Uncertainty Highlighted

The check boxes, recently introduced, serve two purposes, of which only one part is functional. First, they indicate whether the value(s) in that specific area of the table should be included in the uncertainty calculations. By the tabular design, some check boxes will control an entire area or collection of information in a parent-child-like relationship. Consider the example in Figure 42, the second checkbox on the left is for the second analyst (William Tell); however, in this example there is no checkmark next to the analyst's frame. This means that although the inner box is marked, it will not be computed because the outer box is not marked and precludes

inner uncertainties from being included in the computation of uncertainty. This allows the user to turn on/off large blocks of uncertainty, which would be reflected in the accumulative uncertainty chart that is available in the upper right of the window. The second purpose of the checkbox was to cause an update to the local uncertainty graph at the top right of the browser, as mentioned in the previous example. This would show the user, in real-time, the impact that particular information or lack thereof has on the uncertainty associated with the object. Unfortunately, as I mentioned, two problems prevented the functionality of these features. First, I could not add data to the XML document object model (DOM) using JScript and Internet Explorer, which was required to recalculate the uncertainty. Second, the security features of HTML, JScript, JavaScript, and ActiveX prevented file access and therefore precluded saving the file. Although I could toggle the calculative checkmarks, I could not change the value for any calculative and meant that recalculating the data would be unnecessary as it would not change the graph. In addition, if I could add data, I would not be able to save the changes to a file, which meant that the map view and gnomon would not change and was unnecessary. Even the examples that Microsoft provided to demonstrate these capabilities did not work. These problems can be overcome as XML matures and standards evolve or through another programming language.

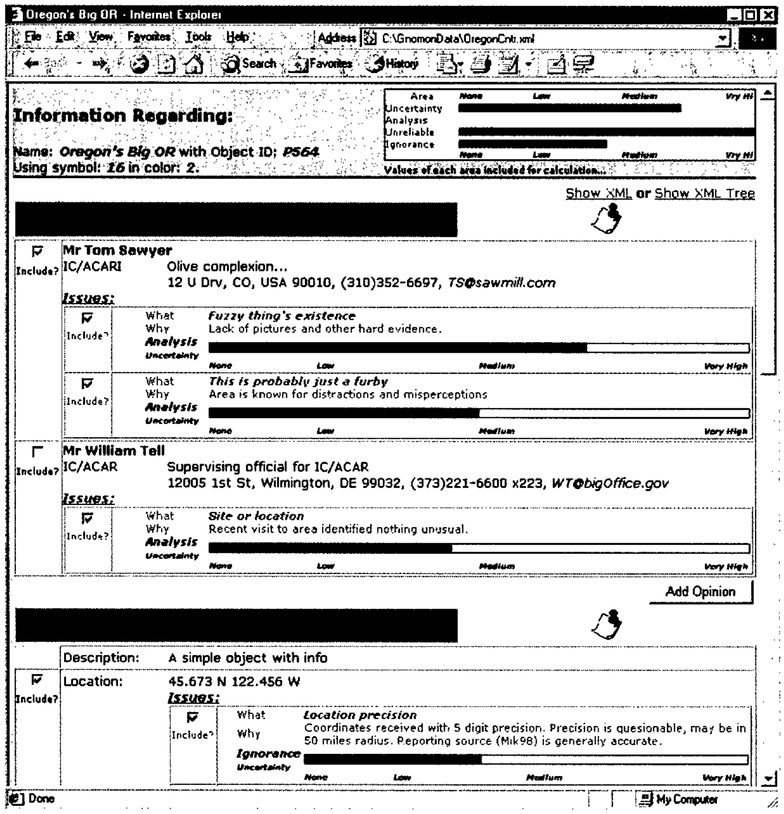


Figure 42. Data Through the Browser

As previously mentioned, the mouse pointer will change in several areas to indicate other information. The post-it notes indicate collapsible areas allowing one to tuck data out of the way. In an attempt to facilitate analytical input, I added a form that is accessible via the *Add Opinion* button, shown on the middle right of Figure 42. This button triggers the opening of another page, shown in Figure 43, that allows the analyst (or authorized person) to add their opinion and rating. Unfortunately, this was only partially functional. Although the form is functional, problems similar to the checkbox were present, namely the inability to add data to the XML DOM and the inability to save the file.

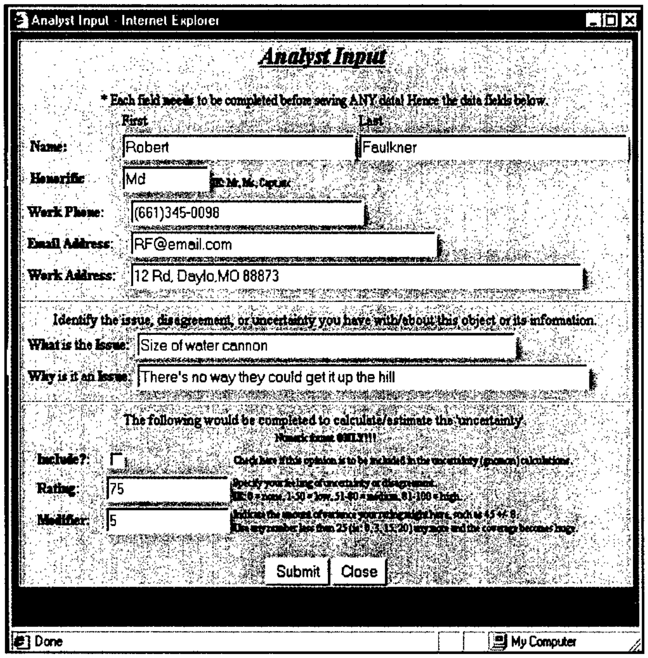


Figure 43. Analyst Input Form

The *Submit Changes* button at the bottom of the main page, shown in Figure 40, also served two purposes. The first was to save the data file, which as mentioned, did not work, and second, to set a trigger that the mapping tool would detect and cause it to reload its data. This function did work and saved a blank file to the data subdirectory using Internet Explorer's ActiveX. Unfortunately, using this technique could confuse users because ActiveX alerts the user to the "possible threat" before the file is saved.

During earlier phases of my research I thought that some elements could be more intuitively expressed using meters or gauges native to or associated with their domain, e.g Figure 44. For instance, fuel levels are usually round with values between E and F, and vertical scales commonly display temperatures. In addition, I thought a more fuzzy logic based implementation could use graphs that more closely represent the measurement function and its uncertainty rather than the simple bars that reflect a scale from low to high. Figure 44 shows a few ideas I was

hoping to implement. One symbol shows there is very little uncertainty in the value while the others show increased uncertainty with respect to their approximate value.

On the other hand, I realized later that these symbols could actually increase the complexity of presenting the data as well as using the information they project. The meters and graphs in Figure 44 may look nice, but they increase the complexity of reading the data. Using any graph in Figure 44 may decrease performance because the user must pause and identify where the item or uncertainty falls, then he needs to determine the highs and lows followed by the uncertainty associated with that information. Finally, the user would have to consider how that specific amount of uncertainty relates to the object as a whole. Conversely, any user can look at a bar graph labeled none–high and figure out what it means. Furthermore, the scale of the bar graph never changes, unlike a representative gauge that would vary by element. For instance, the fuel indicator is easy to read, but the impact of low fuel is only determined by applying that knowledge to that particular object. A jet with low fuel may not be able to complete its mission, while a motorcycle with low fuel could result in a different finding. An extension to this idea would be to include a toggle that allows the user to switch between using easier to read bar graphs like Figure 41 or more representative gauges like Figure 44.

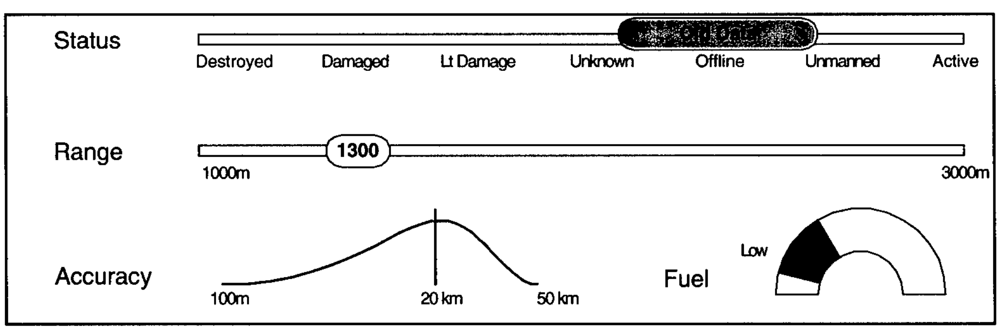


Figure 44. Fuzzy Meters

3.6.4 Areas for Improvement

Provided below is an itemization of issues associated with this prototype that need work or could improve the concept and tool.

1. Map Area and Java API, updating XML node values. The SUN Java XML method used to change or save values to the in-memory DOM tree does not work. The data is currently unchangeable through this interface.
2. Retest gnomon in a more realistic GIS platform with data, roads, boundaries, man made features. Examine the presentation of gnomon fuzz on a map interwoven with other symbols.
3. Attempt to identify the point when gnomon becomes clutter or indecipherable among streets and normal objects then determine an alternative solution.
4. Internet Explorer, JavaScript, JScript, and XML. Resolve XML DOM and file saving issues. Find or create a process to save the XML document from the browser. Other people are trying VBscript, and Java applets.
5. Input. Identify and create a method for creating an object's file and adding to it.
6. Calculating uncertainty. Determine a technique for computing uncertainty that is acceptable to users, then implement it.
7. Graphing of the overall uncertainty and checkbox interaction. Find a way to label checkboxes dynamically. Then recalculate and adjust the graph when a checkbox is toggled. This requires the pre-completion of items 4 and 6. User should be able to enable/disable a check box that in turn changes the calculative attribute of the XML node to true/false and then adjusts the calculations and graphs accordingly. The map could be adjusted or reloaded after saving the modified file.
8. Security and standardization issues also require consideration. Clearly, end users and analysts must learn of any changes. However, a deeper issue involves program management and addressing several configuration management questions: How will data be added to the system? At what point is it added to the knowledge base? What format to use for the data? Allowing analytical input? Processes for updating information to reduce or change uncertainties. Allowing users to maintain local data for various test scenarios. Who is qualified to add and who to remove?

3.7 In Review

To reiterate, the purpose behind this approach, the map and the browser view was to enhance the quality of information the analysts and decision-makers use. The problem that triggered this research was the failure of decision support systems to recognize and handle the uncertainty that exists in the system beyond all attempts to remove it. I believed that we could

improve the analyst and decision-maker's perspective by expressing the uncertainty that was present, and I speculated that we could include other uncertainties so long as we identified them. In turn, the enhanced information would facilitate intuition and decision-making.

This chapter covers two significant areas: the identification of a technique for expressing multi-dimensional uncertainty and an approach for enhancing the analyst and decision-makers information. I explained how the Taxonomy of Uncertainty fostered the identification of quadrants and gnomon as the basis for identifying the uncertainty associated with an object in the graphical display of a DSS. The quadrants reveal the presence of uncertainty that is associated with analytical input, or caused by ignorance and unreliability, which come together into an accumulative uncertainty quadrant.

The uncertainty visualization technique consummated the development of the approach to enhancing the information used in the DSS. The approach involves five processes: (1) recording the uncertainty; (2) identifying objects and elements relevant to the situation at hand; (3) computing the uncertainty based on those elements identified; (4) visualizing the uncertainty; and (5) providing user interaction with the uncertainty.

While evaluating different avenues for addressing the problem of ignoring uncertainty, I found that natural language and fuzzy logic could provide the foundation for this approach. I specified several reasons why natural language is the most appropriate format for recording the information and uncertainty. One reason pointed to the flexibility that fuzzy logic has over precise numbers and the potential for dynamic evaluation based only on the natural language contents of an object's file. The second reason discussed the imprecision of uncertainty, which suggests that a precise method of denoting, computing, and expressing uncertainty is not as crucial as we might think. In fact, the same fuzzy terms and natural language used to retain the data could be used to express the uncertainty.

Although the prototype demonstrates very little use of fuzzy logic, we can employ fuzzy logic in every aspect of this approach. The visualization of uncertainty embodies an approximation through the gnomon, which uses only four lengths to express the uncertainty that is associated with an object. The least involvement of fuzzy logic will occur when a user interacts with the uncertainty by toggling the inclusion or removal of different elements. On the other hand, the user might eventually be able to add his own details that could be composed of natural language.

4. *Evaluation, Results, and Analysis*

By perceiving the enemy and perceiving ourselves; there will be no unforeseen risk in any battle.

Sun Tzu
Art of War

To validate the approach and techniques presented over the last two chapters, I created a prototype that demonstrated the potential employment and viability of these ideas. Fundamentally, the approach facilitates the inclusion of information and objects with uncertainty thereby presenting a more detailed and informed view of a situation or scenario. The program, scenes, and information reveal these benefits by presenting a workspace similar to DIODE that facilitates the inclusion of uncertainty. The demonstration would visualize the uncertainty captured in the data to several NAIC intelligence analysts, showing them the gnomon and the concept behind uncertainty visualization. A supplementary survey completed by evaluators and discussed in this chapter corroborates the concept and utility of including more information by visualizing uncertainty.

The prototype discussed in Chapter 3 exists as a proof of concept and validation of the direction behind this research. This chapter supplies the intent behind the demonstration as well as an explanation of the survey created to evaluate the prototype and its approach. The assessment of the evaluators survey responses makes up a majority of this chapter, which concludes with some final thoughts and inferences.

4.1 Evaluation Intent

The concept of uncertainty visualization and use of gnomon were the focus of the demonstration. The uncertainty visualization facilitates the use of information with uncertainty, which provides a more thorough presentation of the available information. The program

demonstrates the uncertainty visualization through a graphical display using a GIS-mapping tool and filtered presentation of the data. The expected benefit behind uncertainty visualization is the provision of additional and valuable information that improves the analyst or decision-maker's perspective. I expected to gain qualitative rather than statistical results from the straightforward demonstration and survey that would indicate whether the ideas behind the model were attempting worthy functionality.

The script provided to the evaluators guided them through several different screens of the mapping tool. Each view introduced, demonstrated, or compared different aspects of the prototype to DIODE. Although the uncertainty visualization technique using gnomon is applicable across a variety of domains, the program and demonstration were created with the intent of presenting an approach to visualizing the uncertainty of multi-dimensional intelligence data to the sponsor, NAIC. NAIC data spans many databases and involves several facets of the information operations domain, which is diverse as well as laden with uncertainty. In the demonstration, they would step through several different scenes exploring the modeled features that would make the inclusion of uncertain information more practical and functional. The following section explains the path scripted for the evaluators as well as some of the functionality that was included.

4.1.1 Gnomon Demonstration Script

The demonstration script could be printed or referenced and read on-line, it consisted of an enumerated list of steps guiding the user through several different views and features of the program while demonstrating the concept and approach taken. The demonstration had two central portions:

1. The GIS-mapping tool that used Java and GLG to present objects as icons on a map and
2. The browser that was used to present specific information and uncertainty related to an object and its data.

The evaluator started with the mapping portion of the prototype, it was loaded by executing a batch file containing the necessary settings. The first view that is presented shows two objects: one with gnomon and one without (i.e., Figure 38 minus the additional notes). The script and dialog box, on the bottom left of the screen, point out the four distinct quadrants, the four lengths, and that the icon itself is not modified. The evaluator is told about the process that loads each scenario and file, which includes the uncertainty calculation. They are also informed about the object selection process that is simulated by the configuration files. Through background processing, the object's data reveals it location, icon type and color. This should indicate to the users that the program could use dynamically acquired data and information. Nothing, except for the map of the country and its surroundings, is static. The map view provides an eye in the sky perspective, similar to national weather maps. Appendix D contains a few other snapshots of the mapping tool.

4.1.1.1 Introducing the Components of the Mapping Tool

Excluding the preprocessing required to determine which files to load, the main difference between the prototype and DIODE, for the NAIC evaluators, is its inclusion of an object that identifies and expresses the uncertainty it encapsulates. The other differences that are apparent to the user include the small dialog box at the bottom left and different buttons. An underlying consideration was to keep the differences between tools to a minimum, allowing the evaluator to focus on the benefit afforded by the inclusion of uncertainty rather than extraneous nuances. The next two sections walk you through the script the evaluators used, and what they saw or should have seen in the process. Section 4.3 examines the survey responses.

At this point, the evaluator is also informed that I duplicated most of the data about each object; only the values for gnomon lengths, coordinates, and names are different. The data itself is not important to the visualization; the lengths, symbol, color, and uncertainties it recorded were

crucial. By now, the user should have a basic idea about the icons and the data behind the objects, noticing that objects with uncertainty were presented with gnomon fuzz.

The next step of the script briefly describes the various features and tools of the map, which are relatively similar to other basic GIS tools and their implementation of DIODE. The mapping functions include zoom, pan, several "other buttons," and left and right click actions. The "other buttons" identified toggles that enabled or disabled the display of the item. For instance, the *map* button toggled the view of the map, and the *label* button turns the icon labels on and off. One of the other buttons, was the *Gnomon* button that toggled the visualization of uncertainty globally. In addition, if the user left [or right] clicked on an icon, a dialog box appears in the upper right corner of the window. A *Gnomon* button is shown in the dialog box if the item selected is an object with uncertainty. In the prototype, yet unlike DIODE, the connections or links do not have individual data or data files, this was not necessary for the demonstration. The *Gnomon* button in the dialog box toggles the display of gnomon for that specific object, individually. Four other buttons, labeled "Demo: 1, 2, 3, 4" provide the user with an avenue for starting either of four demonstrations or scenes. Right clicking on an object would start Internet Explorer, which loads the object's data as specified by the file link that is passed to the browser from a variable of the icon. At some point, the script suggests the user right-click on an icon with gnomon to examine the information presented through the browser. They can click on several, but as indicated most of the data is replicated and not unique. The browser portion of the script is discussed at the end of this section.

4.1.1.2 Presenting Various Scenes

Following familiarization, the script takes the evaluator through each of the four different views of possible scenarios. The first scene (Demo 1) shows 10 DIODE objects and points out their simplicity and lack of uncertainty. These 10 objects represent what they might see in a

typical system that does not express uncertainty or one requiring very high certainty in the data. The next three scenes show other possibilities that have uncertainty, but are still practical and functional within the visualization environment.

For instance, the second scene (Demo 2) presents a scene with four different objects, as compared to the previous scene. One is without gnomon and implies total certainty while three of them display various amounts of gnomon and indicate the presence of uncertainty. Each of the three icons shows a different amount of gnomon demonstrating different possible combinations. The script guides the users through observing the objects and noting the differences, yet they are similar in size and the presence of uncertainty, they are also legible and distinct.

The third scene shows 16 objects; 12 additional icons were basically added to the previous scene. The addition allows the user to keep the previous icons in mind until the additional icons and their diversity of uncertainty are displayed. Most include uncertainty and few are without gnomon. The view is a lot busier, but should still be legible. Zooming-in clears the view up and the differences between certain and uncertain objects should be apparent.

The final (mapping tool) demonstration presents a worst-case scene: a very busy display presenting certain and uncertain objects composed of the DIODE objects from scene 1 and the most of the objects from scene 3. This is shown in the second figure of Appendix D. By including numerous objects with uncertainty, this view epitomizes a drastically different perspective than those previously presented, and very different from what the analyst would be familiar with or expecting. The ten DIODE objects, without uncertainty, are spread from the west to the north-east parts of the country. Without uncertainty visualization, the southeast part of the country would be barren; however, the southeastern section is covered with icons and gnomon. This should show the user that the mock-country had some defenses that would be omitted and unrecognized if absolute or high certainty were required of the data. If high certainty was required, several objects would be shown, but many others would be omitted. The user should

also note that with the increased clutter, objects are less distinguishable, but it is still clear that there is a lot of uncertainty and information that would not have been displayed in DIODE. Zooming in reduces the close proximity of each icon and clears up the display.

By now, the evaluators have stepped through the predefined scenes that gave them a good idea of what the gnomon could look like and the difference that including uncertainty could make. The script suggests they explore the prototype and its other features.

4.1.1.3 Noticing the Data in the Browser

Returning to a the browser portion, mentioned earlier in this review of the script, the user is guided into right clicking on an object. The right-click invokes Internet Explorer and loads the data file of the object that was selected. Appendix E and Figure 36 through Figure 43 contain several snapshots of the display presented by the browser. Via the script, he is referred to one of the supporting documents describing what they see in the Internet Explorer view and its intended use. The supporting document points out the features discussed in Chapter 3, particularly, the adapted uncertainty visualization used in the browser where a graph of the object's overall uncertainty is presented in the upper right of the browser. In addition, bar graphs depict the uncertainty associated with individual elements in the same location as uncertain data. Several other features are presented as well as some that did not work in the prototype. I would expect the user to be impressed with the niceties as well as the use of colors and contrast making areas with uncertainty information stand out. Compared to the DIODE presentation of the data this view organizes the data differently and draws your attention to the data and uncertainties.

Eventually they conclude their examination of the prototype and different aspects of the implementation and complete a survey.

4.2 Survey

The prototype demonstrates two ways of visualizing uncertainty: gnomon in the DSS and bar graphs in the browser. The survey, a series of 56 questions (in Appendix F), was designed to identify the evaluator's views with respect to the gnomon fuzz visualization, the approach to visualizing uncertainty, the viability of the concept as an enhancement to current tools, and finally, the presentation of data and its specific uncertainty using the dynamic features of the browser. The survey was not designed to comment on issues out of my control, such as the map itself, the mapping program, Java, or Internet Explorer. In addition, a final three-question area wrapped up the evaluation and solicited a final answer as to whether the analyst would use this tool.

In view of the uncommon environment and hectic schedule of the analysts, they required special treatment. I determined that the best way to get the most participation and feedback was to have the sponsor's point of contact handle the distribution of the demonstration and survey. This way the point of contact could provide the demonstration and survey (as proctor) to fellow analysts and evaluators at their convenience as well as respond to questions. The use of the sponsor as proctor was suitable because he was somewhat familiar with the concept, approach, and issues I was working on, plus had witnessed the program in use during a previous meeting. As a testament to his comprehension, he installed and ran the demonstration on another computer.

The respondents would answer each question indicating how much they agreed with each statement by using a number 0–10. Zero (0) indicated that they did not agree at all with the statement while 10 meant they agreed wholeheartedly or completely with the statement. The survey included many related and inter-related statements that I used to crosscheck responses. The survey focused on acquiring a qualitative determination of the following concerns:

- If the gnomon was a clear and distinguishable method of visualizing the uncertainty? If the four lengths of the gnomon were distinctly noticeable? If the quadrants are each uniquely identifiable?

- If the analyst could recall, after a simple and brief introduction, that the different lengths imply various levels of imperfection? Inference: the cognitive load is low, and the gnomon lengths are intuitive.

- If the gnomon or the objects with gnomon are negatively distracting? Did the gnomon interfere with other objects in the mapping tool?

- If the objects with gnomon stand out and are conducive to understanding "what's going on?" If the objects with gnomon impede understanding or the view of "what's going on in the display?" If color would improve their visualization?

- If there is any difficulty with understanding and using the concept of uncertainty visualization by analysts and decision-makers? If the analysts, who regularly work with this information, expect the extra information to be initially complex and a bit of an overload? In addition, would using the gnomon to present the information, be a usable solution?

- Is there potential degradation to the analyst or decision-maker's ability to read a map or reduce their ability to respond in a timely manner?

- Do the analysts believe the gnomon and inclusion of uncertainty can improve the current system? Is there any value added by doing this?

- If analysts add uncertain information to their personal or the entire knowledge base? If they would feel comfortable adding uncertainty without uncertainty visualization to the entire knowledge base?

- If the browser view and its presentation of uncertainty is clear and legible?

- If overall uncertainty graph in the upper right corner of the browser was practical?

- If there is any support for others to have the ability to add information after the data has entered the entire knowledge base. If the analysts would like the ability to add uncertainty information to their data for both personal and global use?

- If the analysts liked the uncertainty visualization? Would they use it?

4.3 Survey Results

Unfortunately, the evaluation did not come off as anticipated, only 5 surveys were completed. Worse yet, only one evaluator witnessed the program in action (the proctor), the remaining evaluators used several snapshots to "see it in action." This sampling was the untimely effect of following an Inspector General's visit (IG), slightly accelerated demonstration

requirements (due to delayed delivery), a snow storm that prevented several members from returning to work, and the lack of an environment with Windows and Internet Explorer. In addition, I could not provide the evaluators with a laptop due to the security restrictions of their environment.

For reasons beyond my control and not for the lack of a functional prototype, the demonstration was reduced to an introduction, a collection of snapshots, and a few pages with the Taxonomy of Uncertainty. Over a week's time, the snapshots and miscellaneous information were distributed to several intelligence analysts of whom 5 surveys were returned. I solicited input from a small class of decision-analysis students due to this poor response to the first survey. The decision-analysis student turnout was also low: two. On the other hand, the two respondents do follow suit with the analysts answers; however, the students' responses tend to be closer to 10 or 0, which implies that the live demonstration and hands-on evaluation provides a better idea of the uncertainty visualization and its functionality. The students' responses (10 & 11) were added to the tabulated responses in Appendix F. Unfortunately, due to the late nature of those evaluations they were not included in the following analysis; however, as mentioned their responses are essentially in concurrence with the analyst responses examined below.

The analysts were described as experienced analysts between 30 and 50 years old, with above average intelligence, good computer experience and an advanced academic degree. A later conversation with the sponsor revealed that most of the analysts were not familiar with uncertainty visualization, the formal implementations and methods of handling uncertainty, nor a comprehensive rationalization or hierarchy of uncertainty. The senior analyst, leading the advanced development of intelligence analyst tools, is however, well versed in the diversity of uncertainty, but not in the hierarchical format or detail the Taxonomy of Uncertainty provided.

The following sub-sections examine and discuss the results of the survey answers tabulated in Appendix F. The following examination of the responses and results to the survey

are neither statistical in nature nor meant to be, they are qualitative, particularly with the small sample size. The survey should confirm the viability of this approach and if the analysts could or would use uncertainty visualization or a similar product in their work.

In the following sections, the term "question" is used interchangeably with the term "statement" and in the context of the survey statements. The survey statements are, in essence, a question soliciting the evaluator's opinion, but framed in statement form for simplicity of the reader, their response, and the evaluation. For instance, the following phrase "questions 2 and 5" refers to the statements 2 and 5. In addition, I refer to the folks that completed the surveys by their survey number (recorded above the results in Appendix F) and the terms evaluator, analyst, intelligence analyst, and respondent.

4.3.1 Gnomon Legibility

Based on questions 1, 2, 3, 6, and 8, all evaluators agree the gnomon and objects with gnomon stand out. The overall higher values show that gnomon is clearly distinguishable method of visualizing uncertainty that does not really interfere with other objects in a GIS decision support tool. The evaluators also noticed that objects with gnomon stood out more than the gnomon-less objects. This is beneficial in the sense that objects with uncertainty will not be missed, but could also be distracting in the long run. User's could be drawn to and distracted by the objects with gnomon, which presents the need for some way to disable the presentation of the gnomon. Two buttons were available in the prototype to toggle the display of gnomon individually and globally; however, this could lead to mistakes by including uncertainty and turning the display of gnomon off. Through 13 and 14, the evaluators specify that the quadrants and lines in each are clearly and uniquely identifiable and distinguishable from the other quadrants. Statement 31 supports their impression that the uncertainty visualization will not interfere with reading the background map or display.

4.3.2 *Uncertainty Connection*

Through question 15, most evaluators admitted they recognized the connection between the gnomon and uncertainty. However, responses by respondent 2 and 5 were so low that it is possible they misunderstood the question, unfortunately neither added a comment to explain their choice. It is possible that the question was misinterpreted in a way that they thought the question was inferring that there was an imperfection in the icon rather than the object it represented. A follow-up interview could also discern if they simply did not understand the correlation between the gnomon and uncertainty. Through further examination, it is apparent that respondent 5 has difficulties and unexpected answers throughout the evaluation. This could be attributed to the lack of a live demonstration and the fact that neither the proctor nor myself were present to clarify any questions.

4.3.3 *Gnomon Characteristics*

High responses to questions 3, 4, 5, and 16 indicate the four lengths of the gnomon were noticeable, distinguishable, and indicated the various levels of imperfection. However, one person (# 2) had some difficulty with the gnomon, its lengths, and presentation; thankfully, he/she included comments between the questions with respect to their troubles. The evaluator commented on questions 16 that it would be "hard to differentiate" the lengths without a reference, which shows that he/she used a snapshot as a demonstration of the program. This is unfortunate because the program had such a reference in the lower left of the main window indicating the quadrants and lengths, similar to Chapter 3 Figure 38, see Appendix D for the actual use. Other than the one person, responses were high.

Through questions 17, 18, 19, 20, 21, and 22, almost all responses clearly indicate that neither longer nor thicker lines would improve the visualization and could actually decrease the legibility of the display. Evaluator 2 indicates, to question 22 that longer lines might help. A

follow-on question, 23, to the line length and width indicates mixed opinion regarding the use of color in the lines. Respondent 2 suggests a mixed approach to drawing the lines using dots, dashes, and solid lines to help denote the difference. A notable suggestion considered in earlier alternatives, but these seemed like they would be either confusing as background objects or noise, or indistinguishable from solid lines because of scaling. On the other hand, the different ways of drawing lines does offer the opportunity to aid the distinction between lengths. Another suggestion, by examiner 2, indicated the use of different colors that contrasted with the background. The use of multiple colors was also considered as a possible solution; however, it was regarded as confusing because of the other meanings that could be applied with the use of color. Furthermore, the different backgrounds (e.g. city, mountains, water) would require alternating line colors in order to maintain the requisite contrast, which clearly increases confusion. Determining the background and contrasting colors of each object would also increase the complexity of the computations behind the visualization.

4.3.4 Gnomon Color

Contrary to expectations, the responses to statements 23 and 24 indicated the lack of an obvious preference or benefit to using color. The mixed results, were more against the use of color than for it. There is no clear indication from their responses that the use of color would improve visualization or their understanding of the information.

4.3.5 Gnomon Interference to Visualization

Questions 7, 8, 9, 10, 11, 12, 19, 20, 21 and 22 convey that the objects with gnomon stand out and are conducive to understanding "what's going on." Most evaluators indicated that the gnomon and objects with gnomon are not negatively distracting, they do not interfere with other objects, and do not impede understanding or the view of situation or scenario. In their

opinion, the objects positively contribute to the visualization. In addition, changes to the line length and thickness would not increase the distinction and could impair the visualization.

4.3.6 Uncertainty Visualization Concept

Statements 25, 26, and 28 indicate the concept of uncertainty visualization was easy to comprehend and would be easy for other analysts and decision-makers to understand. The results of question 27 indicate that very few intelligence (or information) analysts and decision-makers would have difficulties with the concept of uncertainty visualization. Interestingly, these results appear to be supported in this survey. One analyst (# 5), who is not the senior analyst or the proctor, expressed some difficulty with the concept and felt most intelligence analysts and decision-makers would not follow or understand the notion behind uncertainty visualization. This insight comes from the same analyst indicating difficulty with the visualization and the indication of uncertainty through the presence of gnomon. Interactive use of the prototype in addition to a complete introduction to the prototype demonstration might have avoided these difficulties. Training is clearly essential to any new concept and product.

4.3.7 Using Uncertainty Visualization

By the responses to 27 and 29, most analysts agree that it would not be difficult for intelligence analysts and decision-makers to use or work with the uncertainty visualization presented in the prototype. On the other hand, the evaluators indicate by statement 30 that uncertainty visualization would affect their ability to respond in a *timely* manner. Comments mentioned the added complexity and data overload as negative impacts to using the included uncertainty when *timely* responses are required. One analyst (# 4) suggested a dual implementation where the analysts, not the decision-makers, see and use the uncertainty. The same analyst suggests that the decision-makers have a view without uncertainty for two reasons:

to reduce the decision-makers complexity and to prevent their reduced confidence in their staff and intelligence analysts by the impression of waffling and incompetence.

In contrast to its use in a time-constrained situation (of 27 & 29), responses to questions 45 and 46 indicate that they do not think the uncertainty visualization would overload the analysts. The analysts reveal that the uncertainty visualization and the complexity of dealing with extra information might not be as much of an issue when used for analysis and less time sensitive work, yet an overload for time sensitive issues. An idea and desire mentioned by the lead analyst was to include a slider or knob that when increased or decreased caused the uncertainty visualized to change accordingly. He envisions a tool that shows objects with uncertainty at or below the level specified by the slider, sort of like a dimmer-light switch increasing and decreasing the amount of light put out. The slider essentially acts as a filter toggling the display of icons when they qualified to be included in the display.

4.3.8 Uncertainty Visualization Improving Accuracy

The analysts' responses to questions 32, 33, 34, and 35, are difficult to lean in a particular direction. However, the middle ranging answers seem to suggest that they are not sure about uncertainty visualization improving their task accuracy due to the additional information or not. Their answers are not decisive in one way or the other; they expect the accuracy will be affected positively as well as negatively sometimes. The use of accuracy was intentionally vague allowing the readers to fill in their own interpretation or sense of accuracy in their work, unfortunately this could have resulted in the inconclusive responses too.

4.3.9 Adding Uncertainty Without Uncertainty Visualization

The responses to questions 36-42 required a bit of interpretation, which is identified here before discussing the results. Two analysts (# 2 & 5) responded abnormally to the follow-on question that asks at what approximate point or percentage do they add objects or information

with uncertainty. I contacted the sponsor for insight to their odd responses because I was expecting low values instead of the high ones I received. Respondent 2 specifies 100% on four questions as the level of uncertainty an object/information has when he adds it to the system. The corrected interpretation of the results is that "100% of the objects and information have some uncertainty, therefore, his responses to 37, 39, 40, and 41 are ineffective. This perspective also applies to question 39 for respondent 1 who also indicated 100%, I know for a fact that #1 is fully aware of the uncertainty present in the system. The responses by respondent 5 were interpretable based upon the responses to 40 and 41, which ask about them adding uncertainty if it were visualized. Examining 5's responses, without interpretation, indicate that he/she would add less uncertainty to the system if it were visualized. Therefore, I expect that he/she meant to specify the amount they were certain rather than uncertain, in which case 36 and 37 in conjunction with 40 and 41 make sense as the uncertainty increases.

The following analysis follows with the understanding presented in the previous paragraph. By 36 and 38, most analysts are not comfortable adding uncertain information without uncertainty visualization to their working data for personal work or research nor the entire knowledge base for others to use. I anticipated these results, but they are more interesting after learning that only one of their numerous databases does not allow for the inclusion of uncertainty. The analysts use databases that require they specify a certainty value between 1 and 5 (5 is the lowest certainty) before the data is accepted and saved. Their results are interesting because this implies some of the analysts may not realize the value of the certainty factor they are using.

Oddly, while previous responses seem to indicate that the analysts are not aware that they are already using an uncertainty indicator the results to 37 and 39 reflect otherwise. By the responses, it seems that that all of the analysts do in fact add (or realize they add) information with uncertainty to their own as well as the entire knowledge base. Respondent 3 seems most

aware of the uncertainty that he/she adds by their response of 10% to the follow-on question. According to the results of 36 and 38, they must add the uncertain information with great discomfort. From the two that indicated they do add objects and information with uncertainty to the their personal data, their uncertainty values differed drastically. By the responses, one analyst (# 5) includes objects and information with approximately 25% uncertainty while the more conservative (# 3) only does so when uncertainty is about 10%.

One analyst's (# 5) response indicates great support for the inclusion of uncertainty; the additional information provided to 40 and difference with 36 indicates he/she would include more uncertainty if it were visualized. Furthermore, by 40, 41, and 42, most analysts would definitely be comfortable adding uncertain information to their own and the entire knowledge base if uncertainty was expressed or identified, respondent 5 would even do so with 50% uncertainty.

4.3.10 Value-Added with Uncertainty and Uncertainty Visualization

Questions 43 and 44 indicate a consensus that there is value added to the DIODE-like tools by including uncertainty and uncertainty visualization. Analysts and decision-makers can benefit from the additional information even thought it includes some uncertainties.

4.3.11 Uncertainty Visualization in the Browser

Responses to 47, 49, and 50, indicate the analysts felt the uncertainty visualization presented in Internet Explorer was noticeable, clear and the approximate values distinguishable. Unfortunately, in retrospect question 48 is poorly worded, they would have seen, in the demo, that any element of any object or information could have many associated uncertainties; the answers should be higher. According to the responses, most did not understand that the elements containing uncertainty and shown in the browser could be attributed to multiple causes. In addition, the analysts agree by question 51 that the overall uncertainty graph (in the upper right of

the browser view) is very constructive as a reference to the entire uncertainty associated with the data.

4.3.12 Adding Information and Opinions

Most analysts support the ability for authorized analysts to add information to the data after it has entered the entire knowledge base, via question 52. On the other, hand one user (# 5) is dead set against it, but wants ability himself to add information as indicated by question 56. In late conversation I had with the lead analyst, he caveats the ability to add comments with the understanding that standards, controls, and data assurance would be addressed before any thing like this occurred. Furthermore, by questions 54, 55, and 56, all of the analysts would like the ability to add uncertain information to their data for both personal and global use. In addition, they would also like the ability to add their opinion and issues to objects and information.

4.3.13 Prototype Overall

As for an overall impression of the prototype and uncertainty visualization approach, question 53, the analysts thought it was OK and needed some work. However, most would like to use it or something like it for their own work.

4.4 Survey Final Thoughts and Inferences

Overall the prototype and approach presented had a positive impact on the evaluators; their responses and the comments by the lead analyst support the potential practicality of uncertainty visualization in their environment. The primary concern and effort is the development of a tool-set for extracting, classifying, correlating, and otherwise intelligently processing data for integration into the knowledge base. Although some degree of uncertainty is being captured in the preliminary stages of the intelligence processing, it is quickly replaced with a number. An idea instigated by the approach I demonstrated was to keep the various causes and

representations of uncertainty throughout the lifecycle of the information rather than masking or replacing them with a single certainty value (1–5).

Taken as a whole, the evaluators reflected understanding of uncertainty visualization and support for the concept and its use in their work. In addition, they indicated that the concept could be slightly challenging for others to grasp and use. On the other hand, they also indicated that uncertainty visualization could improve their work and the work of others as well as degrade timeliness potentially through its complexity. The results also indicate an overall discomfort with using or adding objects and information to the knowledge base with uncertainty without the use of uncertainty visualization; even though, they are already doing it. Furthermore, all would like the ability to add their opinions and issues to information and objects

In terms of visualization, the consensus indicated that the present approach is easy to see and notice; the lines should not be colored, widened nor lengthened. Furthermore, the quadrants and lines in each were distinguishable and the lengths generally notable. Overall, they could identify the different amounts of gnomon between objects; some had more gnomon than others did. The evaluators liked the use of the cumulative graph used in the browser and indicated that the uncertainty areas were obvious and bar graphs clearly indicated the approximate level of uncertainty.

The respondents provided several interesting and feasible suggestions. One alternative for uncertainty visualization was to constrain the uncertainty visualization to the analytical environment to avoid presenting the uncertainty to a decision-maker. The reasons for not presenting the decision-maker with uncertainty were to reduce the decision-makers complexity and to preclude loss of confidence in their staff. I disagree with this suggestion and believe the decision-maker should see the uncertainty involved in their decision-making environment; the key to including uncertainty is the improved knowledge gained by understanding what is uncertain and what is not. On the other hand, I do realize that some commanders may not want to

see or use the uncertainties and some situations will call for the exclusion or minimization of uncertainty. However, this should not preclude the availability of the information, whether or not it is used.

A viable suggestion that facilitates the inclusion of uncertainty without its visualization was the use of a slider or knob, mentioned earlier, to adjust the inclusion of uncertainty. The adjustment to the visualization added or removed objects as the slider/knob was changed. The idea has been a desire of the lead analyst for a while now. His ideal environment would provide a slider that increases or decreases the objects on screen by increasing the allowed uncertainty in the environment. This is viable in the sense that pre-selection or filtering determines the objects for inclusion in a scenario. In this simple prototype, the objects below a certain threshold of uncertainty could be enabled and visualized while those above are disabled. Then as the knob is turned, the threshold adjusted, and objects within that were not being presented, displayed or otherwise displayed because they exceeded the threshold could be re-evaluated for visualization. The slider/knob essentially acts as a filter toggling the display of icons when they qualify to be included in the display.

Although the demonstration and evaluation of the prototype was not as thorough as I would have preferred, I believe the program and techniques encompassed for organizing data, visualizing uncertainty, and presenting the data are successful. The strategy behind their development followed guidelines for information visualization as well as uncertainty visualization to ensure uncertainty was recognized and expressed clearly. In addition, the responses indicate an awareness of the uncertainty and the ability to distinguish its parts when not obscured by many icons.

5. Findings and Conclusion

Belief systems at rest tend to stay at rest unless a force acts upon it, much like the laws of conservation in the physical world.

Rob Lambert

One of our greatest and frequently accomplished challenges is change. Every day some thing is improved and our routine changes; however, change is often difficult and unwelcome, which is what I foresee for the eventual inclusion of uncertainty in military decision support systems. A paradigm shift will occur when we start visualizing uncertainty in decision support systems in which command decisions affect the fate of soldiers, but I expect we will see these changes within the next five to ten years.

5.1 In Review

The thesis began by elaborating on the common occurrence of information overload bolstered by our information technologies and the all too frequent inundation with extraneous minutiae. Information is knowledge, yet too much is useless and burdening unless presented in a manner that makes translation and inception practical. In addition, some of that information involves uncertainty that must be handled appropriately.

Undeniably uncertainty exists everywhere and in everything, particularly throughout information and data employed in military decisions. One shortfall of the current decision support systems is the failure to include uncertainty as well as its exclusion of objects and information that are uncertain or have uncertainty associated with them. My research aspiration was to find an approach to enhance the information the analyst and decision-maker used by including uncertainty and information with uncertainty. However, if uncertainty is used in a system, then the users should be alerted to its presence and use.

Uncertainty visualization is a scheme of supplementing the principal information with information about the uncertainty thereby providing a more comprehensive and accurate presentation of information for users to analyze. Unlike some visualization applications, the uncertainty present in military DSSs involves numerous classes of data and therefore multiple dimensions of uncertainty. Early awareness of the complexities involved with multi-dimensional uncertainty indicated the requirement for an all-purpose dimension-free approach to visualizing uncertainty.

One aspect of defining a solution to visualizing multi-dimensional uncertainty involved the classification of uncertainty, which effectively led to my establishing the Taxonomy of Uncertainty. The Taxonomy of Uncertainty transpired from numerous resources, classifying two fundamental categories of uncertainty followed by four types of ignorance and three types of unreliability followed by many sources of uncertainty. The taxonomy assisted in the development of the uncertainty visualization strategy by inspiring the use of quadrants to express the uncertainty that was associated with any object.

I partitioned an icon into quadrants then augmented it with gnomon, which provided a technique for indicating the two fundamental classes of uncertainty. The multiple dimensions of uncertainty that could possibly be associated with an object were expressed through two quadrants and presented in combination with a third dimension, the analytical input, in a third quadrant. In conjunction with intelligence analysts' desires to include issues and uncertainties of their own with an object or information, the quadrants designations unfolded to represent analytical input, cumulative uncertainty, unreliability, and ignorance.

In the end, a prototype was created to demonstrate the approach devised over the course of studying other work involving cognition, heuristics, reasoning, uncertainty, and information visualization. The prototype embodied the idea of identifying the elements that are used in the calculation of uncertainty through its use of the calculative attribute. The program also computed

and used the uncertainty recorded in the data; it then presented the uncertainty using gnomon whose lengths were specified by the calculated uncertainties. In addition, the prototype provided a sense of interaction with the data by presenting the data via the browser and including checkboxes that the user could (conceptually) toggle to adjust the computations of uncertainty.

The Taxonomy of Uncertainty that helped identify a technique for visualizing uncertainty can be used to identify and classify uncertainty in data and information. I presented the taxonomy, concept, and prototype to the National Air Intelligence Center, the organization sponsoring the research, for exploratory considerations and evaluation.

5.2 Findings

The prototype's implementation demonstrated the concept and viability of my approach for enhancing the information provided to analysts and decision-makers in DSS. It also showed that the uncertainty visualization employing quadrants and gnomon successfully expressed information in a useful manner that indicated to evaluating analysts the presence of uncertainty.

One of the issues behind including uncertainty in a DSS is the necessity for identifying and recording the uncertainty of any given scenario. Although originally amassed and defined in the shadows of respected cognitive scientists to determine a sound approach to visualizing multi-dimensional uncertainty, the Taxonomy of Uncertainty also facilitates the identification of uncertainty and causes for uncertainty. The taxonomy provides a detailed, hierarchical means of identifying and classifying uncertainty. The initiating architect and designer of DIODE conveyed that the concept and approach to the uncertainty visualization was insightful. In addition, he noted that the taxonomy is practical for the discovery and classification of uncertainty in their intelligence gathering tools.

Although the approach presented in this thesis is discussed with respect to a DSS, the approach can be applied across other domains and activities, namely through DIODE's potential

to support logistical and political endeavors. In general, this approach to handling uncertainty by representing multiple dimensions of it appears to be extensible to many scenarios.

Finally, and referring back to Chapter 1, the OODA loop is improved by enriching the knowledge on which observations and decisions are made. Through uncertainty visualization, the analyst and decision-maker are provided with a perspective that allows them to more accurately identify and understand a situation and adjust their orientation accordingly. The expected results are better decisions and refined actions.

5.3 Future Research and Direction

It seems like more time and energy went into learning about the issues that go into designing a successful system than went into the creation of the prototype. Looking back on this, I realize that most of the work was to validate earlier ideas I had about a potential implementation. The actual coding of the prototype occurred late in the thesis development, but was reasonably clear because I had resolved several issues that could have posed a problem. By the time I started coding, I had a scheme for visualizing the multi-dimensional uncertainty. The tough part was getting the program and XML to do what they were supposed to be able to do. With the abundance of information explored and presented, I can envision several other uses for this research. I discussed many issues and alternatives in the material providing a number of opportunities for follow-on research. Several if not most are reiterated in the following sections.

5.3.1 Extending the Taxonomy of Uncertainty

The resultant Taxonomy of Uncertainty is inspiring and survived criticism by the half dozen AFIT faculty and students that it was provided to for comments. The critics included professors from the following areas Information Resource Management, Decision Analysis,

Artificial Intelligence, Information Warfare, and Mathematics and Neural Networks. The other two examiners included a doctoral and a graduate student.

The taxonomy's hierarchy and functionality should be validated against a variety of intelligence sources. Although assembled in conjunction with correlation to previous categorizations of uncertainty, the taxonomy was not applied against information to identify and classify the uncertainty identified in that information. One might apply it against several data sources in the accumulation of information for one object then compare it against the uncertainty identified by skilled and unskilled analysts compiling data. A different consideration is the implementation of the taxonomy in an automated process for filtering data. However, a more comprehensive collection of terms is required for every reason and cause for uncertainty before this attempt. Its extension could certainly encompass a catalog of subordinate and parallel terms identifiable as causes to those already established as leading to uncertainty. This collection of terms would be incomplete since represent a countable infinite set of words. The taxonomy and catalog could be shared with any number of intelligence gathering and information processing organizations.

5.3.2 Improve and Validate on Real GIS

Formalize this approach for including uncertainty, including its reference to the quadrants and gnomon. Most systems have a standard scaling factor for their symbols; the gnomon could be defined as a ratio of that scale enabling dynamic length and width calibration.

The plain background of the prototype makes for a simple implementation. In addition, the Java-GLG extension proved to be troublesome; I recommend using another programming language or design environment to reproduce the approach presented in a realistic GIS.

It may even be possible to develop uncertainty visualization as a portable component or package with specific IO parameters. This would have many applications; JOVE might be a practical target. Section 3.6.4 includes some other areas for improvement.

Another idea is the development of an implementation where each quadrant line of ignorance and unreliability, as mentioned in Section 3.4.4, represents one of values for the six types of uncertainty. Visualizing the individual lines could be functional in other domains.

Finally, it could be very useful to include the uncertainty slider suggested by an NAIC analyst, the creator of DIODE, in Section 4.3.7. Having a slider or dimmer that controls the display of objects based upon the amount of uncertainty that is associated with them, as the dimmer is increased the regulated allowable uncertainty is also increased and vice versa.

5.3.3 Fuzzy Computation of Uncertainty

Another opportunity exists in the fuzzy logic realm. The prototype described herein uses a weak technique for computing the uncertainty associated with an object; exploration and refinement of some of the included ideas could be extended to virtually any domain recording uncertainty. Section 3.3 Techniques for Computing Uncertainty provides some information. A refined approach to combining or maintaining distinction between different domains is also exploitable.

5.3.4 Visualizing Multidimensional Data in 3D

Several military systems are moving into the 3D realm, for instance JOVE is expected to eventually be a DSS for commanders. Can this approach be extended into the 3D domain? One might determine whether the approach for uncertainty visualization of multi-dimensional data is applicable to 3D visualizations.

5.3.5 Uncertainty in Textual Systems

Other than the graphical systems discussed in this research, there are text-based systems that may require some way of representing uncertainty. The development of a text-based method for expressing uncertainty could be examined. An idea is to apply various formatting to the text to imply different meanings. Some possibilities are mentioned in 3.4.2.2 Textual Considerations.

5.4 Final Thoughts

In retrospect, this thesis exploded into an immense research effort requiring an extensively diverse background, or lots of work – I worked very hard. I hope the material and references I include can be useful to someone else so they may avoid half the work. On the other hand, it also resulted in the consumption of a plethora of knowledge and new respect for decision analysis, HCI, fuzzy logicians, and information visualization.

5.5 Conclusion

This thesis identifies and explains an approach to enhancing the information used by analysts and decision-makers without unduly increasing their burden. It is accomplished by including and handling uncertainty in the decision support system. The inclusion of uncertainty is made practical by techniques for visualizing the uncertainty that is present; the strategy involved gnomon and bar graphs that successfully expressed the presence of as well as an approximation of uncertainty.

The complete approach includes five major areas for development. They consist of recording uncertainty, selecting the objects and elements for inclusion, computation and visualization of uncertainty and finally, the interaction with the uncertainty.

In a broad and risky generalization, I contend that we will never be free of uncertainty and as such, I suggest that a refined and formalized version of this approach be incorporated into

future decision support systems. We must recognize that uncertainty is prevalent and we should provide for the explicit inclusion and presentation of uncertainty to alert the user of its presence so he can use it to his or her advantage or eliminate it from inclusion.

Appendix A: Complete Taxonomy of Uncertainty

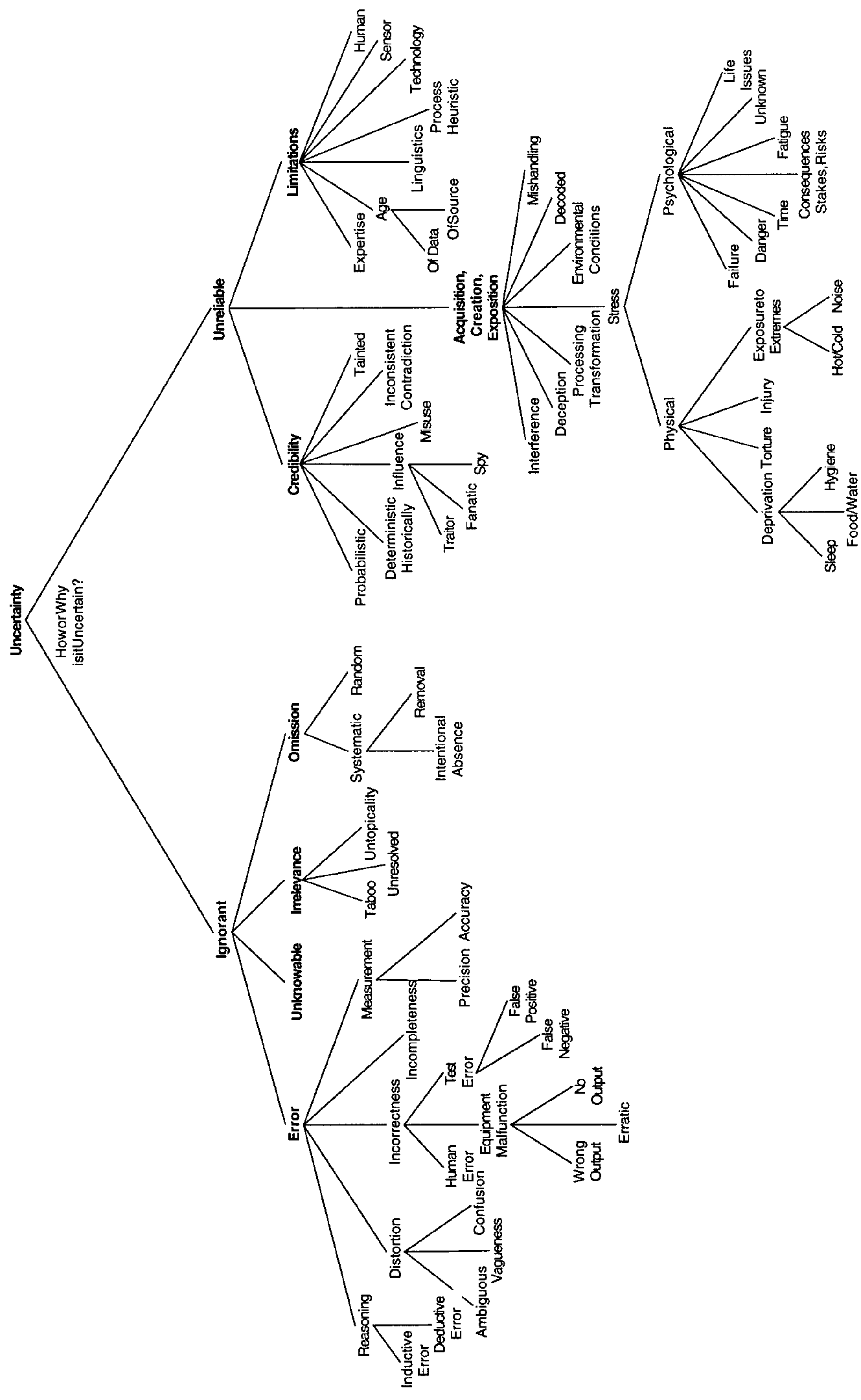


Figure 45. Fully Expanded Taxonomy of Uncertainty

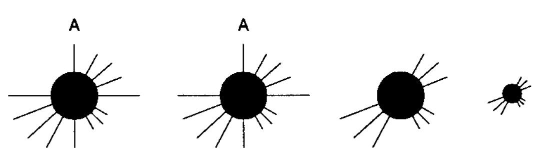
Appendix B: Comparing Gnomon Fuzz Organization.

Obviously, there are many combinations of rays, lines, and fuzz that could be evaluated but only a few provide distinction worth evaluating. I evaluated the following objects by means of a simple comparative examination to identify the best use of lines to represent uncertainty. The samples also demonstrate additional confusion that could be added while trying to improve the decision-makers information.

There are five sets of objects labeled A through E. For the sake of consistency, most object lines (excluding D) are three simple and identifiable sizes: small, medium, and long. The lines of A, B, C, and E are placed in the center of the quadrant and equidistant from each other. The lines of D would be spread equally and vacant areas, where attributes are not shown, are not removed to allow for equal dispersion. This prevents indicator sliding, where a line moves left or right just because an attribute is or is not shown; maintains location consistency.

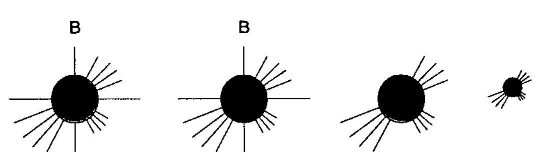
Three features were considered: symbol clarity, distinction of lines or quadrants including the smaller size, and recognition of four line lengths (none, short, medium, long) except for D.

- **A** represents an object where the uncertainty is divided into quadrants and uses three lines to represent the level of uncertainty in each quadrant.



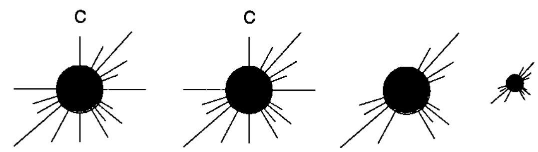
- Observation: The quadrants are recognizable, clear and uncluttered. In addition, the lengths are notable. Use of three lines is simple but looks a little too spaced apart.

- **B** represents an object where the uncertainty is divided into quadrants and uses four lines to represent the level of uncertainty in each quadrant.



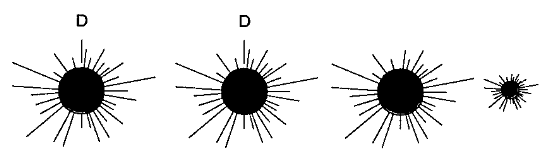
- Observation: The quadrants are recognizable, clear and uncluttered. Line lengths are notable. The four lines appear to stand out more than three lines. Seems like it would support color more than the 3 lines. The blending of the four lines would aid the perception of the color, as long as only one color was used in each quadrant.

- **C** represents an object where the uncertainty is divided into quadrants and uses four distinct lines in each quadrant. Each line in each quadrant represents the uncertainty for one or more different attributes.



- Observation: The quadrants are less obvious. The mixed lines throw the perceptions off. The long diagonal lines force the viewer to twist their head a little trying to align their natural visual horizon with the diagonal. This forces and incorrect visualization of the object. It would be difficult to discern what the four different lines in each quadrant meant. Even more so for the uncertainties of the smaller object where lines seem to melt together.

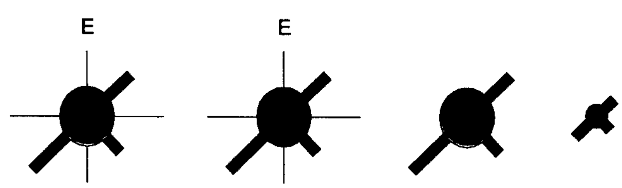
- **D** represents an object where the uncertainty is marked by lines coming out of every angle. There are NO distinct quadrants, each line could represent the uncertainty of a given attribute.



- Observation: The multiple lines make recalling a specific attribute very difficult. It is difficult to identify one line from another. The lines mix and blend too much for practical individual use. The smaller object starts to look like a splat or spill, as it becomes a fuzz ball.

  * On the other hand, these lines or fuzz could capture the uncertainty of many attributes and would be useful when observing the changes of all of the fuzz. Knowing each line would require too much training. It would be better to notice that many lines change when one or more fields were changed or a particular idea/scenario was changed. It could also be useful to notice how all of the fuzz changes based upon the addition or removal of some input(s).

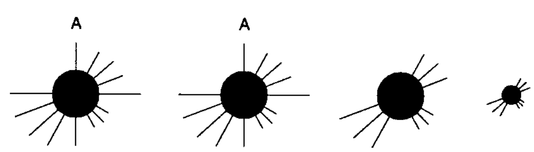
- **E**, like A, B, and C represents an object where the uncertainty is divided into quadrants, except that this object set uses one thick line to represent the uncertainty in each quadrant.



- Observation: The quadrants and lines are recognizable, clear and uncluttered. Each line stands out but also changes the object. Clearly the thick bar would support color however the size of the lines take focus from the object itself and become new parts of the object. The smaller object looks like an entirely new symbol, which is clearly an undesirable side effect.

It is easy to see the number of options available for depicting information around an object. It is also easy to simply conclude that adding a line for every attribute is desirable since it shows all or as much of the information available; demonstrated by object D. However, by observing object D we also see that the complexity of reading the lines increases with the number of lines. The far left image of Object D, above, was magnified 150% to accurately and easily count the number of lines – 29. 300% magnification is required to discern the lines of the smallest in set D. Remember, the purpose of visualization is to present some information in a manner that induces further thought, not to overload the viewer with more information that requires additional digging.

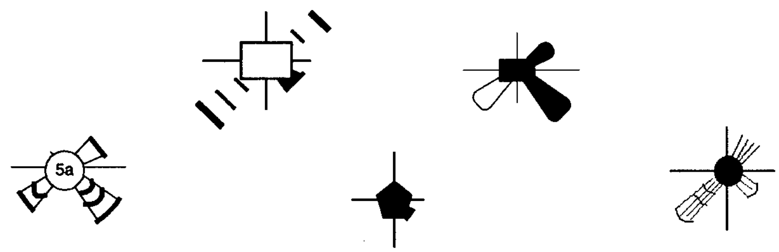
The clearest depiction of information stands out through objects A and B using three or four lines in each quadrant. Of the two, object A is slightly better. Objects in A use three lines that are close enough they even enhance the viewing of that quadrant without almost changing the look of the object. Four lines, as in B, seem to change the shape to look almost like a hard candy. Using three equidistant, same-length, lines per quadrant will be more visible, easily read and less confusing.



Other Ideas

Some other ideas and improvements considered but thrown out because they failed to improve the visualization of uncertainty are shown below. One idea incorporated color to express

additional warning or another dimension of uncertainty. A second idea used lines as crosshatches to more clearly indicate the level exceeded by the lines or area. The third idea used enlarged areas to either stand out more or express another dimension.



Appendix C: Object Symbols Used.

The following numbered icons were included in the GLG palette and used by the program. Icons 12-18 come from MIL-STD 2525. Below are a few basic icons from DIODE and MIL-STD 2525.

Icon Number	Description
0	Basic icon from map demo; blue circle w/ label & data area
1	Small circle, filled in
2	Plus symbol (+)
3	X symbol
4	X with a dash through it
5	Open circle
6	Open square
7	Filled square
8	Open diamond
9	Open diamond with a plus symbol (+)
10	Open star
11	Open triangle
12	Open square with X i.e. infantry
13	Diamond with double tipped arrow and bar across: SnglRocktLnchrLite-Red
14	Diamond with double tipped arrow and 2 sides: MltiplRocktLnchr-Red
15	Diamond with radar dish and lightning mark: GrndTrkEqpSensrRadar-Red
16	Diamond with missile and bar across: AirDefShrtRng-Red
17	Diamond with line coming off mound and H across: AirDefGunLite-Red
18	Diamond with line coming off mound and triple bar-H across: AirDefGunHvy-Red

Icon 4 

Icon 6 

Icon 8 

Icon 10 

Icon 13 

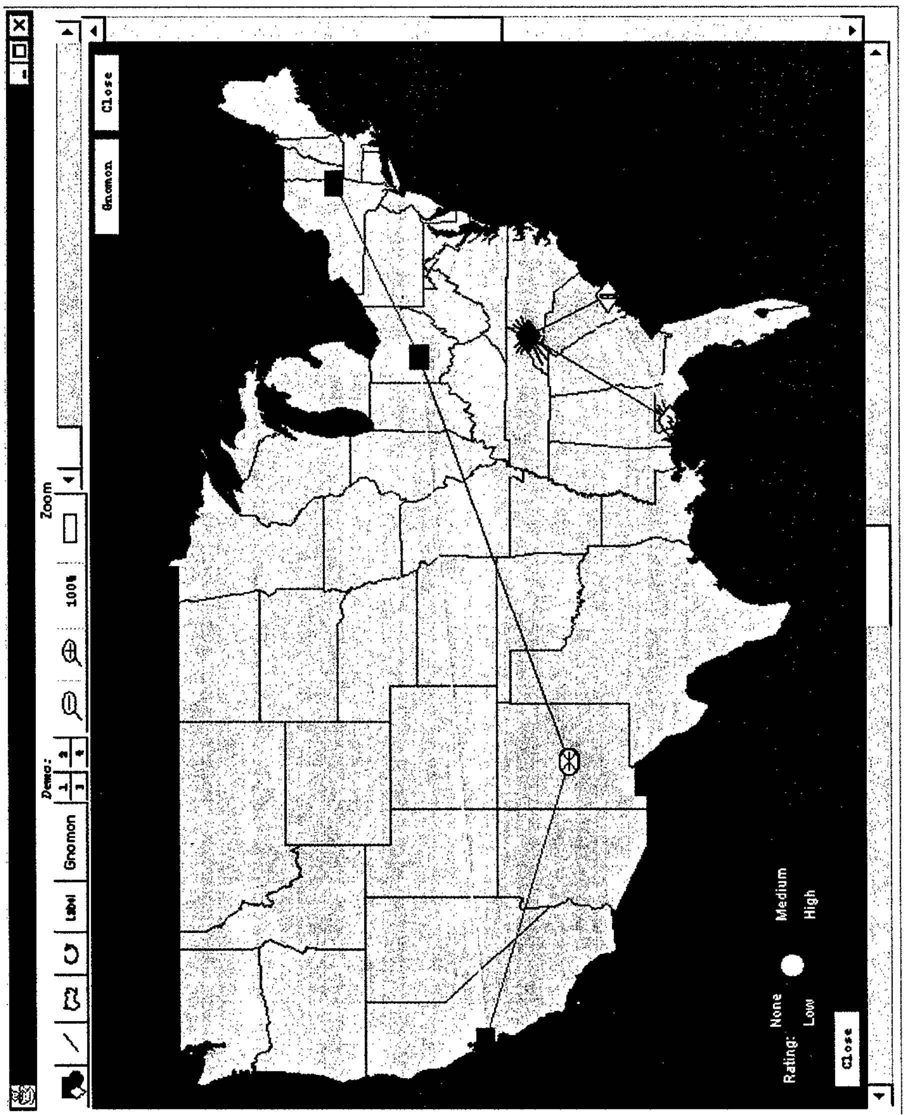
Icon 15 

Icon 16 

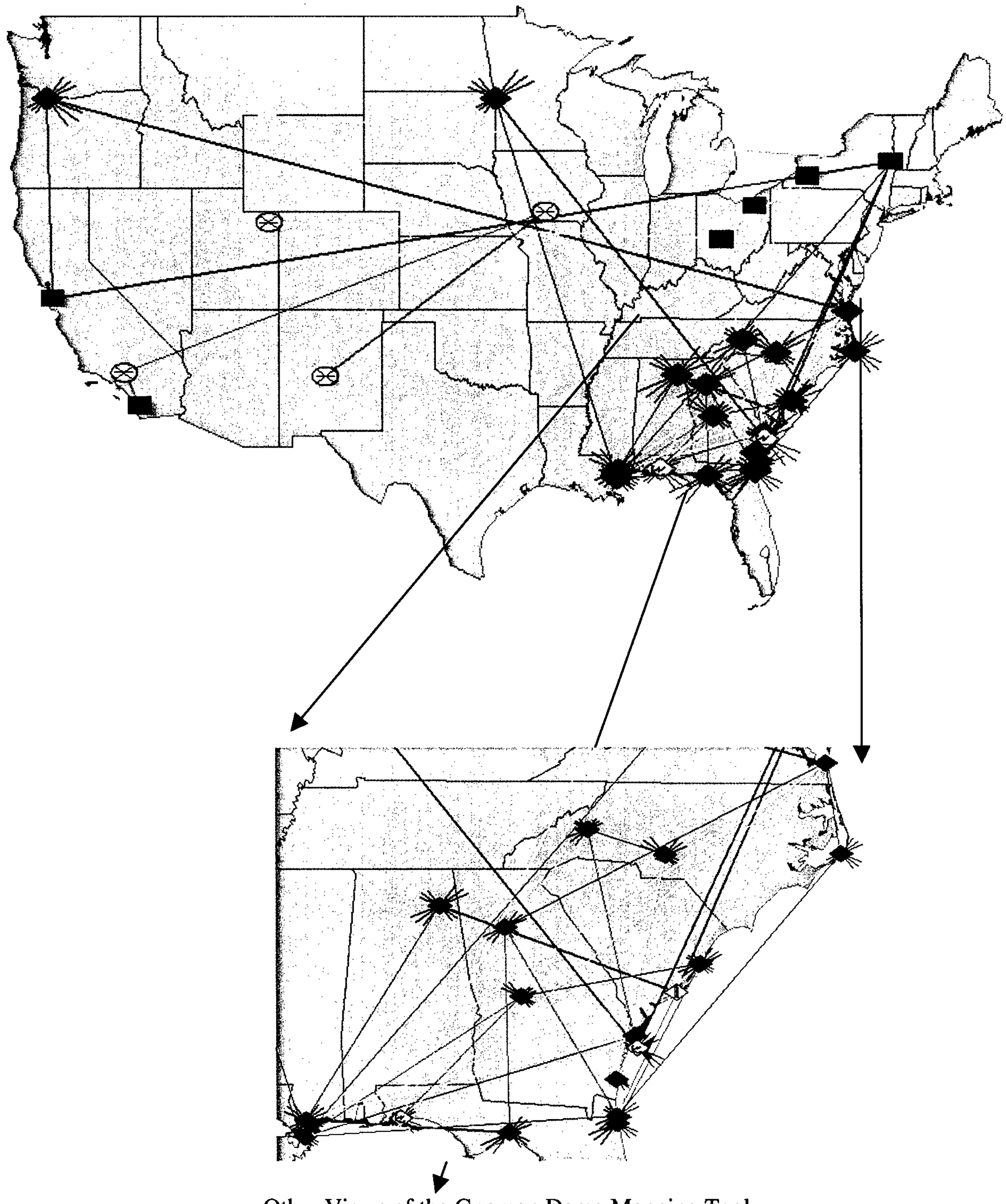
Icon 18 

Appendix D: Prototype Mapping Tool Snapshot.

The following image is a snapshot of the mapping tool demonstrating the use of gnomon indicating the presence of uncertainty. The plain blue boxes and white circle represent symbols used in DIODE with no uncertainty while the others, in the southeast, taken from MIL-STD 2525 are augmented with gnomon fuzz.



Gnomon Demo Mapping Tool

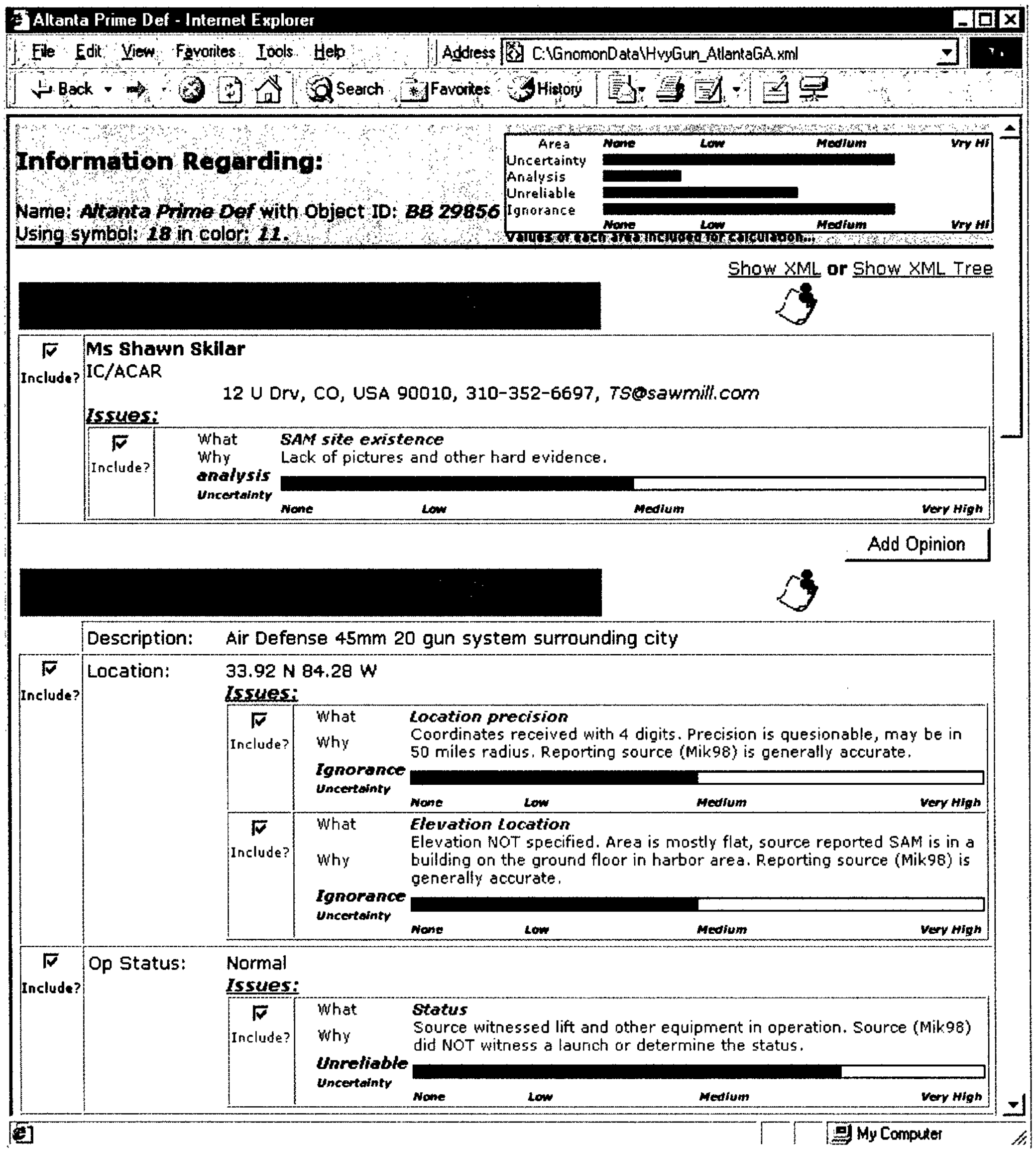


Other Views of the Gnomon Demo Mapping Tool

Appendix E: Prototype Presentation of the Data.

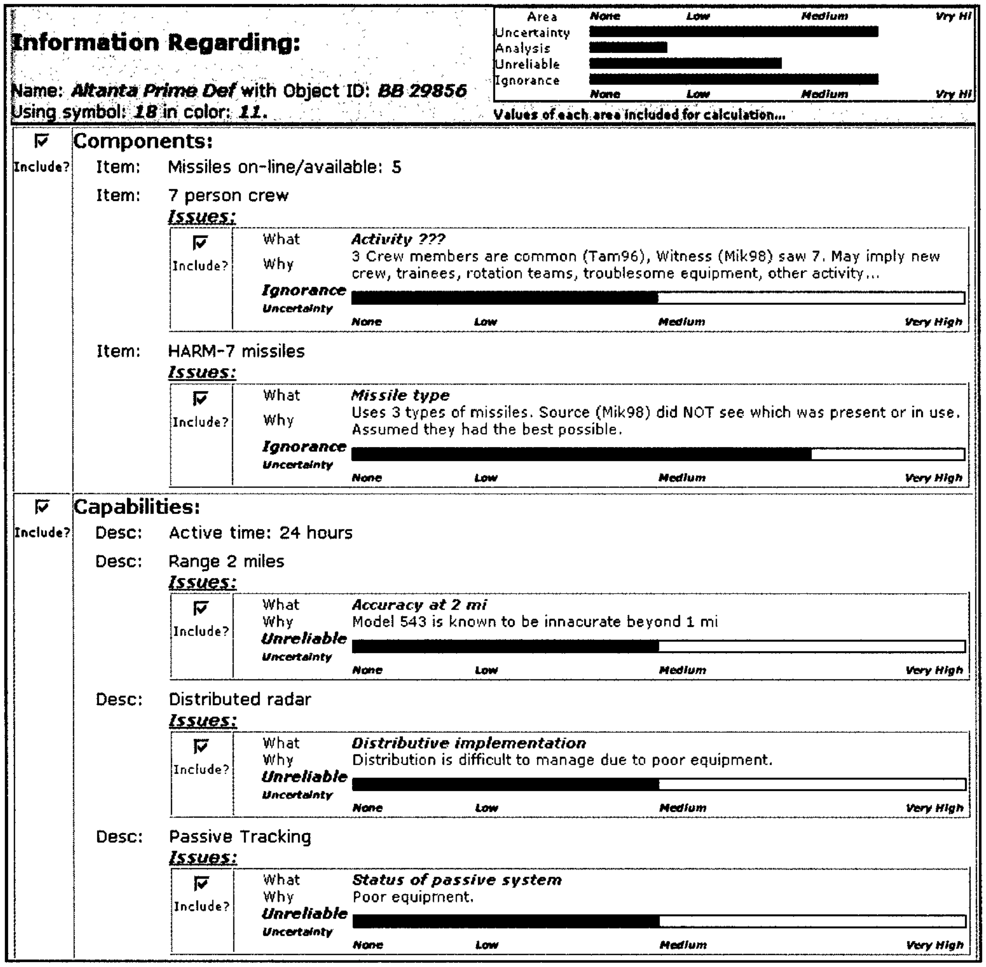
The snapshots demonstrate the formatting techniques used to present the data for any given object represented in the mapping environment. Most importantly, it ensures the uncertainty present in a file stands out.

The use of different colored tables, shown in the figure below, draws the viewer's attention and the bar graph summarizes the amount of uncertainty associated with each particular element.



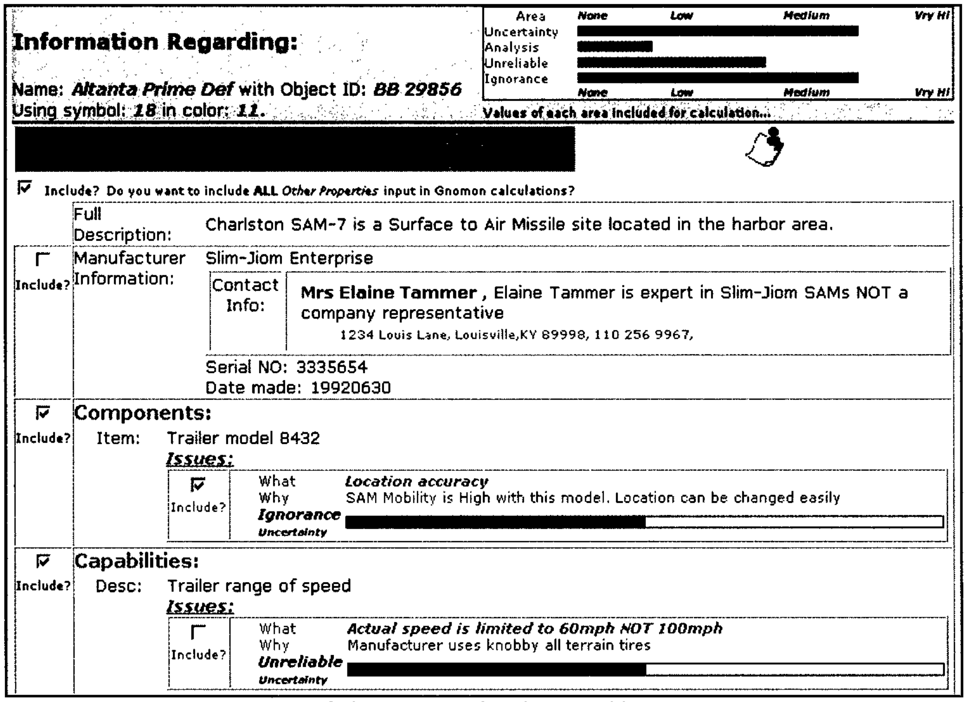
Initial Internet Explorer View

This figure shows a screenshot of other information included in the file for the "Altanta Prime Def" object. The misspelling of Atlanta is intentional and infers human involvement and continuing potential for error no matter how small or large. The components and capabilities shown below are part of the Key Properties starting in the previous figure on the preceding page.

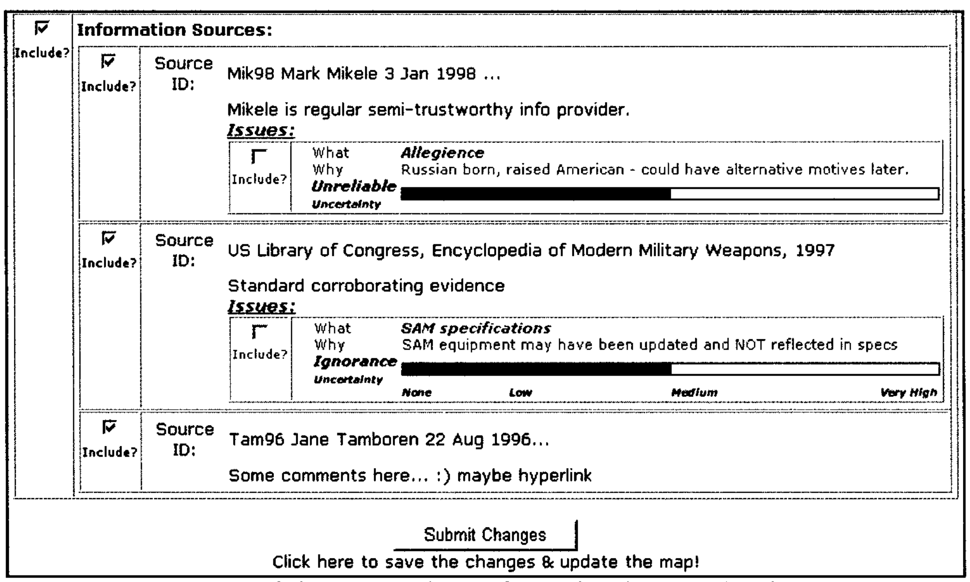


Key Properties, Components & Capabilities

The next two figures show part of the content of the Other properties section in this file. Notice that the graph in the upper right presenting overall uncertainty remains visible.



Other Properties Screen Shot



Other Properties' Information Sources Section

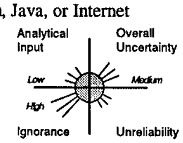
Appendix F: Prototype Demonstration Survey

First, Thank you for your time, assistance, and patience, your participation is greatly appreciated. Second, your personal information is not required however, **please note that this survey is subject to review**. Comments or suggestions are used to suggest improvements and future study but may be reviewed by the NAIC sponsor as a security measure and the thesis committee for verification.

The purpose behind this map and the browser view is to improve the quality and amount of information the analyst and decision-makers are provided in decision support systems like DIODE. The problem behind this research is that we fail to include objects or information in our decision or problem space when we are uncertain about them or some portion of their information. The theory is that we can include the information and objects with uncertainty if we also identify those issues. This prototype is one way to do that. Using gnomon (fuzz) like on the nearby circle we can include various imperfections while indicating the presence of an object/information that might have otherwise been omitted.

Constraints: Colors and Object symbols are provided/specified by standards and CANNOT be changed/replaced. Colors vary by affiliation in conflict: friendly, unknown, foe, neutral. ∴ Must add to or augment the symbol.

The prototype is one possible way of visualizing uncertainty. The survey, a series of statements, is designed to identify your views on the gnomon fuzz visualization, this visualization of uncertainty, and the viability of the concept as an enhancement to current tools. The survey is not designed to critique the US map, the mapping program, Java, or Internet Explorer. The three sections are



- **Regarding the gnomon and its visualization.**
- **Regarding the concept of visualizing uncertainty.**
- **Regarding the browser view of "detailed" information and specific uncertainty.**

And finally, a three-question area wraps up the evaluation.

* Yes, some statements are reworded to support previous answers, not to trick you.
** **Emboldening** is used to emphasize key words and ensure distinction is noted between previous or similar statements.

Please indicate how much **you agree with** each statement. For each statement specify a number between 0 and 10, see below for a sample value line. In view of the fact that many people wear many hats and fill many rolls, you are welcome to complete additional surveys if you are answering them from different perspectives (e.g. first as analyst then as a decision-maker). Alternately, you can use different colored pens for different roles. In addition, each section provides space for any suggestions and comments.

|---|---|

0	5	10
Not At All	Somewhat	Completely, 100%

Please identify the roll or perspective from which you are completing the survey:
Analyst   User, non decision-maker   Decision-maker

<u>Regarding the **gnomon fuzz** and its visualization.</u>

1. The objects with gnomon* fuzz are distinguishable from other objects without fuzz. _____
2. I can tell which objects have uncertainty associated with them because of the gnomon. _____
3. It is easy to identify those objects that have more gnomon fuzz than others. _____
4. I can make simple approximations (none, low, medium, high) about the amount of uncertainty associated with the objects that have gnomon fuzz. _____
5. Of the objects with gnomon fuzz, I can identify those with more gnomon fuzz than others. _____
6. The objects with gnomon stand out among those without fuzz (e.g. Demo 4). _____
7. The objects with gnomon are **not** distracting. _____
8. The objects with gnomon **do not** interfere with the identification of normal objects. _____
9. The objects with gnomon **do not** interfere with the identification of any other objects. _____
10. I expect the gnomon to interfere with my view of "what's going on" in a real display or map. _____
11. The objects with gnomon fit into the map without standing out negatively. _____
12. The objects with gnomon contribute positively to the visualization. _____
13. I can easily identify the four quadrants. _____
14. The lines in each quadrant are distinguishable from the other quadrants. _____
15. The presence of gnomon indicates a level of imperfection and uncertainty to me. _____
16. The four lengths (none, low, medium, long/high) of gnomon fuzz are distinguishable. _____
17. Thicker lines would improve the visibility of the gnomon. _____
18. Thicker lines would make the four lengths more distinguishable. _____
19. Thicker lines would interfere with the map and other objects. _____
20. Thicker lines would decrease the readability of the map and background objects (if other data were on the map). _____
21. Longer lines would not improve the visibility of the gnomon. _____
22. Longer lines would interfere with the map and other objects. _____
23. Different colored lines would improve the visibility of the gnomon. _____
    What color would you suggest? _______
24. Multicolored lines would improve the visibility of the gnomon fuzz. (e.g. black for low, yellow for medium, red for high). _____

Suggestions / Comments:
Please provide any suggestions and comments you might have in the following areas:

- Gnomon fuzz width, length, color:
- Use of quadrants to display gnomon:
- Use of quadrants to breakdown the uncertainty:

Regarding the concept of visualizing uncertainty.

We use uncertainty visualization to increase understanding of what we are using or looking at by increasing our knowledge of what is uncertain. In many cases, this is accomplished by including information that might otherwise be omitted due to its imperfections. Often, by visualizing the variances or possible values it is possible to express the uncertainties. The uncertainty visualization in the prototype presents objects that have uncertainty associated with them. Gnomon fuzz indicated and approximated the presence of the uncertainty related to various objects.

25. The idea of visualizing the uncertainty presented in the introduction and used in the demo was easy to comprehend. _____

26. The concept of uncertainty visualization would **not** be **difficult** for most analysts and decision-makers to understand. _____

27. It would be difficult for very few analysts and decision-makers to **use** or **work with** the concept of visualizing uncertainty. _____

28. This use of uncertainty visualization (using gnomon fuzz) would be **difficult** for most analysts and decision-makers to understand. _____

29. It would be **difficult** for some analysts and decision-makers to **use** or **work with** the uncertainty visualization presented in this demo. _____

30. Uncertainty visualization would **degrade** the ability of most analysts and decision-makers to respond to or "interpret a scenario" in a timely manner. _____
Why? _______________

31. Uncertainty visualization would **degrade** the ability of most analysts and decision-makers to read a map or display. _____

32. Uncertainty visualization would **improve** the **accuracy** (for target selection, event prediction, foe evaluation, etc) of most analysts and decision-makers. _____

33. Uncertainty visualization could significantly **improve** the **accuracy** of analysts and decision-makers in some areas. _____

34. Uncertainty visualization would **decrease** the **accuracy** of analyst and decision-maker views, analysis, or predictions (e.g. through overload). _____

35. Uncertainty visualization would **decrease** the **accuracy** of analysts and decision-makers very little or not at all. _____

36-41: The terms "knowledge base" refers to any collection of information (databases, notes, or records) used to retain data and information about an event or object. For instance, the "entire knowledge base" refers to one of the many master databases used by end users after information is validated. Personal "knowledge bases" represents collections of information that may not be completely validated but is used to aid further work in related areas.

36. I am comfortable **adding** objects/information with uncertainty to **my knowledge base or environment** for personal work/research/analysis **without** using uncertainty visualization. _____
If you are comfortable, then at what approximate point or percentage do you add them? _____

37. I currently **add** objects/information **with uncertainty** to **my knowledge base** for **personal** work/research/analysis. _____
If you do, then at what approximate point or percentage do you add them? _____

38. I would **NOT** be comfortable **adding** objects/information with uncertainty to the **entire knowledge base** for **others to use** in their work/research/analysis **without** using **uncertainty visualization**. _____
If you are comfortable, then at what approximate point would you add them? _____

39. I currently do NOT **add** objects/information **with uncertainty** to the **entire knowledge base or environment** for **others to use** in their work/research/analysis. _____
If you do, then at what approximate point or percentage do you add them? _____

40. I would be comfortable **adding** objects/information with uncertainty to **my knowledge base or environment** for **personal** work/research/analysis **if** uncertainty were expressed in some way (e.g. gnomon fuzz). _____
If you would be comfortable, then at what approximate point would you add them? _____

41. I would be comfortable **adding** objects/information with uncertainty to the **entire knowledge base or environment** for **others** to use in their work/research/analysis **if** uncertainty was expressed in some way (e.g. gnomon fuzz). _____
If you do, then at what approximate point or percentage do you add them? _____

42. I would NOT be comfortable **using** objects/information with uncertainty **if** the uncertainty **were not** expressed in some way like through the gnomon fuzz. _____

43. There is value added to visualizing uncertainty in DIODE-like tools. _____

44. There is value added by including objects with less than perfect certainty in the analyst and decision-maker knowledge base. _____

45. Uncertainty visualization would overload the analyst and decision-maker initially with a more useless information. _____

46. Uncertainty visualization makes the display of information too difficult and complex to use. _____

Suggestions / Comments:
Please provide any suggestions and comments you might have in the following areas:

- Uncertainty visualization for the analyst:

- Uncertainty visualization for the decision-maker:

- Increasing the accuracy of analysis:

- Reducing the time to interpret or read a display that includes uncertainty:

<u>Regarding the browser view of "detailed" information and specific uncertainty.</u>

The presentation of an object's information via the browser was an extension of the uncertainty visualization, not the focus. This view presented specific information related to the object and any uncertainty associated with it.

47. The uncertainty areas (titled **Issues** and lighter in color) stood out noticeably. ____

48. It was clear that any area of uncertainty could have many different causes or sources resulting in uncertainty. ____

49. Each bar graph clearly indicated the amount of uncertainty associated with a particular item or piece of information. ____

50. The overall uncertainty graph area (on the top right) was **clear.** ____

51. The overall uncertainty graph area (on the top right) was a **useful** way to look at that object's uncertainty and its causes for uncertainty. ____

52. I support the ability for authorized personnel/analysts to add information and **their** opinion and issues to any object and its overall uncertainty **from the browser after** it has been entered into the entire knowledge base. ____

<u>Suggestions / Comments:</u>
Please provide any suggestions and comments you might have in the following areas:

- Presentation of the uncertain information related to an object:

- Presentation of the information associated with an object:

- Uncertainty representation through bar graphs, dial gauges or meters, sweeps, etc.:

- Use of an area representing an object's overall uncertainty:

<u>Final Questions</u>

53. How would you rate this visualization of uncertainty using gnomon fuzz (circle choice)?
    1. Useless; there is NO way I would add ANY uncertain information to the knowledge base.
    2. Terrible; everything about this is counter-intuitive or against common sense.
    3. Not too bad; idea needs work, visualization needs some work, AND I'm not comfortable including uncertainty.
    4. OK; idea is understandable, visualization needs some work.
    5. Neat; gnomon fuzz needs improvement, I need more guidelines before I would use it (e.g. what and when to add info).
    6. Good; I would use it IF ____________________.
    7. Great; I would use it for personal analysis ONLY.
    8. Excellent; I'd use it today if it were implemented AND if everyone was familiar with it.

54. I would like the ability to include uncertain information in my knowledge base for personal use. ____

55. I would like the ability to include uncertain information in the entire knowledge base for everyone to use. ____

56. I would like the ability as an authorized analyst to add **my own** opinion and issues to the object and its overall uncertainty. ____

<u>Suggestions / Comments:</u>
Please provide any suggestions and comments you might have in the following areas:

- Including uncertainty in a DIODE-like environment:

- Uncertainty visualization:

- Anything else...

Survey Results

*Empty fields denote that no answer was provided. Double horizontal lines delineate survey sections. Boxed cells indicate possible misinterpretations, which are discussed in the analysis. Surveys responses 1–5 came from NAIC analysts, responses 10–11 came from the decision analysis students.

Table 7. Survey Results

Q #	Survey Responses 1	2	3	4	5	10	11	Avg
1	8	5	8	8	7	10	10	7.2
2	7	7	9	9	10	10	10	8.4
3	7	6	10	9	6	9	10	7.6
4	8	6	9	8	5	9	8	7.2
5	8	8	7	9	6	9	8	7.6
6	7	6	7	9	7	10	10	7.2
7	8	6	8	4	7	10	10	6.6
8	8		9	8	8	10	10	8.3
9	8		10	8	8	10	10	8.5
10	3	5	2	4	3	9	2	3.4
11	8	6	9	8	6	10	9	7.4
12	8	6	9	8	5	10	9	7.2
13	10	7	10	10	7	10	10	8.8
14	10	8	9	10	7	10	10	8.8
15	7	3	10	10	2	10	8	6.4
16	8	3	7	7	6	8	6	6.2
17	2	2	3	2	2	3	1	2.2
18	2	2	3	2	2	3	1	2.2
19	8	8	8	8	9	8	9	8.2
20	3	8	8	9	9	8	8	7.4
21	8	4	8	8	9	3	8	7.4
22	7	7	8	8	9	10	9	7.8
23	2	8	5	6	2	7	9	4.6
		3 contrasting ones	black					
24	2	8	6	2	2	7	2	4.0
25	9	8	9	9	4	9	8	7.8
26	8	8	5	9	3	9	10	6.6
27	7	5	5	6	9	3	10	6.4
28	2	3	3	3	8	3	2	3.8
29	2	3	7	2	8	3	2	4.4
30	7	7	7	5	5	0	1	6.2
	data overload	complexity	overload					

Q #	Survey Responses 1	2	3	4	5	10	11	Avg
31	3	4	3	3	3	3	1	3.2
32	7	3	7	1	6	7	9	4.8
33	5	5	7	3	6	8	9	5.2
34	5	4	6	5	3	3	1	4.6
35	5		7	8	7	7	1	6.8
36	5	3	2	2	8	0	10	4.0
					$\boxed{75}$		$\boxed{95}$	
37	9	10	7	10	10	5	10	9.2
		$\boxed{100}$	10		$\boxed{75}$		$\boxed{95}$	
38	5	7	8	9	10	6	3	7.8
39	1	3	8	0	0	8	3	2.4
	$\boxed{100}$	$\boxed{100}$		uncert				
40	8	10	6	10	8	5	7	8.4
		$\boxed{100}$			50			
41	8	10	6	9	8	8	7	8.2
		$\boxed{100}$			50			
42	8	6	7	9	10	10	2	8.0
43	9	7	9	6	9	10	10	8.0
44	6	7	9	9	9	9	9	8.0
45	3	3	6	9	2	3	1	4.6
46	3	4	4	5	3	4	1	3.8
47	5	8	5	9	9	9	8	7.2
48	7	4	5	3	6	9	8	5.0
49	8	7	8	9	7	10	9	7.8
50	8	6	10	6	7	9	9	7.4
51	8	8	10	6	9	8	8	8.2
52	10	7	9	10	0	10	7	7.2
53	4	5	7	7	5	7	8	5.6
54	10	10	9	10	9	8	8	9.6
55	10	8	6	10	8	8	8	8.4
56	10	8	8	10	8	8	8	8.8

Bibliography

[Ack98] Ackerman, Robert K. "New Display Advances Brighten Situational Awareness Picture." *Signal.* AFCEA Publications, Washington DC, MD, August 1998.

[Ack99] -----. "Visualizing Information Emerges as Major Element of Operations." *Signal.* AFCEA Publications, Washington DC, MD, 1999.

[AHD98] American Heritage Dictionary of the English Language, Third Edition. Houghton Mifflin Co, Boston, MA, 1996.

[AVC99] Advanced Visualization Technology Center. "Computational/Measurement Errors." Http:// www. avtc.org/error.htm, accessed 28 January 2000.

[AW99] Agosta, John M, and Weiss, Jonathan J. "Sources of Uncertainty". *Diagnostic Uncertainty.* Knowledge Industries Inc, http:// www. kic.com/id24_m.htm, accessed 8 October 1999.

[Bob99] Bobo, Allen. "Thesis Exploration Interview." With respect to DIODE, working environment, and some analyst issues. NAIC/TAIT, Wright Patterson AFB, OH, May 1999.

[Cha97] Jeong, Chang-Sung and Pang Alex T. "Uncertainty Visualization of Hierarchical Information Structure Using Reconfigurable Disc Trees" University of California Santa Cruz, Santa Cruz, CA, 1997.

[Cle96] Clemens, Robert T. Making Hard Decisions An Introduction to Decision Analysis. Duxbury Press, New York, NY, 1996.

[CMS99] Card, Stuart K, Mackindlay, Jock D, and Shneiderman, Ben. *Readings in Information Visualization Using Vision to Think.* Morgan Kaufmann Publishing, San Francisco, CA, 1999.

[CMU93] Carnegie Mellon University "What is Fuzzy Logic" *Fuzzy FAQ.* School of Computer Science, Carnegie Mellon University, Pittsburgh PA, 15 April 1993.

[Cox99] Cox, Earl. *The Fuzzy Systems Handbook, Second Edition.* Academic Press, New York, NY, 1999.

[DA95] Department of the Army. *Military Operations: Battlefield Visualization Concept.* TRADOC Pamphlet 525-70. Headquarters, United States Army, Training and Doctrine Command, Fort Monroe, VA, 1 October 1995.

[Dah96] Dahl, Arden B. *Command Dysfunction: Minding the Cognitive War.* Master's thesis, Air University, School of Advanced Airpower Studies, Maxwell AFB, AL, February 1996.

[DG00] Defense Information Systems Agency "Global Command and Control System Interoperability Home Page." DISA/GCCS IOP, Washington DC, MD, http://jitc.fhu.disa. mil/gccsiop/ accessed 31 January 2000.

[DII98] Department of Defense. *Defense Information Infrastructure Master Plan*, Version 7. DISA, Washington DC, MD, 13 March 1998

[DIO99] National Air Intelligence Center. "Introduction to DIODE." Wright Patterson AFB, OH, November 1999.

[Edw82] Edwards, Ward. "Conservatism in Human Information Processing." *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York, NY, 1982.

[Fad95] Fadok, David S Maj. *John Boyd and John Warden: Air Power's Quest for Strategic Paralysis.* Masters Thesis, Air University, School of Advanced Airpower Studies, Maxwell AFB, AL, February 1995.

[Fis82a] Fischhoff, Baruch. "Debiasing." *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York, NY, 1982.

[Fis82b] -----. "Those Condemned to Study the Past." *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York, NY, 1982.

[Fle93] Flemming, Richard W. *Vulnerability Assessment Using a Fuzzy Logic Based Method.* Masters Thesis, Air University, Air Force Institute of Technology, Wright Patterson AFB, OH, December 1993.

[Fog99] Fogg, B.J. "Persuasive Technologies" *Communications of the ACM.* Vol 24 No 5. ACM Publishing, New York, NY, May 1999.

[GLG99] Generic Logic Inc. *GLG Toolkit.* Boston, MA, www.genlogic.com, 3 October 1999.

[GM88] Gulick, Roy M and Martin, Anne W. "Managing Uncertainty in Intelligence Data - An Intelligence Imperative." *Science of Command & Control; Coping with Uncertainty.* AFCEA International Press, Washington DC, MD, 1988.

[GR94] Giarratano, Joseph and Riley, Gary. *Expert Systems, Principles and Programming.* PWS Publishing Co, Boston, MA, second edition, 1994.

[Har97] Hardebeck, Daniel J. "Television facts." *The Kill Your TV Website*, 1997. Http://othello.localaccess.com/hardebeck/killtv2.htm, accessed 28 June 1999.

[Hoc97] Hockenbury, Don H and Hockenbury, Sandra E. *Psychology.* Worth Publishing, New York, NY, 1997.

[Hom00] Homer, James H. Discussion about DIODE, prototype demonstration, survey, and more. Homer is credited as the initiator and designer of DIODE. NAIC, Wright Patterson AFB, OH, 26 January 2000.

[Kel93] Keller, Peter R and Keller, Mary M. *Visual Cues. Practical Data Visualization.* IEEE Computer Society Press, Piscataway, NJ, 1993.

[Kor97] Korte, Tom. "JOVE Brings 3D Battlespace to Life" *InterCom.* Air Force Communications Agency, Scott AFB, IL, May 97.

[KT81] Kahneman, Daniel and Tversky, Amos. "Variants of Uncertainty" *Cognition.* Elsevier Sequoia, New York, NY, 1981.

[KT82] -----. "Judgement Under Uncertainty" *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York, NY, 1982.

[Lea99] The Learning Kingdom, Inc. Palo Alto, CA, http:/_/www.LearningKingdom.com, accessed 22 October 1999.

[Lev97] Levkowitz, Haim. Color Theory and Modeling for Computer Graphics, Visualization and Multimedia Applications. Kluwer Academic Publications, Boston, MA, 1997.

[Mat89] Matlin, Margaret W. C*ognition.* Holt, Rinhart, and Winston Publishing, Fort Worth, TX, 1989.

[McK98] McKenna, Pat. "What's the Big Idea?" *Airman.* AF New Agency Kelly AFB, TX, October 1998.

[MDB87] McCormick, BH; Defanti, TA; and Brown, MD. *Visualization in Scientific Computing.* Report to the NSF Advisory Panel on Graphics, Image Processing and Workstations, August 1987.

[MF93] McNeill, Daniel, Freiberger, Paul. *Fuzzy Logic.* Simon and Schuster Publishing, New York, NY, 1993.

[Mil56] Miller, George A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity For Processing Information." *The Psychological Review.* American Psychological Association Journals, New York, NY, March 1956.

[MFJ96] Miller, Gregory J LtCol; Fecht, Kurt C Maj; Jefts, Barbara Maj et al. *Virtual Integrated Planning and Executions Resource System (VIPERS): The High Ground of 2025.* Air Force 2025 Future study presented to Joint Chiefs of Staff, November 1996.

[MSB98] Microsoft Bookshelf 1998. *Quotations.* Pliny The Elder (c. 23–79), Roman scholar.

[Ord99] Orden, Karl F, Nugent, William, LaFleur, Bonnie, and Moncho, Silvina. *Assessment of Variable Coded Symbology Using Visual Search Performance and Eye Fixation Measures.* Naval Health Research Center, San Diego, CA, 12 January 1999.

[Osb96] Osborne, William B LtCol (USA); Bethel, Scott A Maj; et al. *Information Operations. A New War-Fighting Capability.* Air Force 2025 Future study presented to Joint Chiefs of Staff, November 1996.

[Osk82] Oskamp, Stuart. "Overconfidence in Case-Study Judgements." *Judgement Under Uncertainty: Heuristics and biases.* Cambridge University Press, New York, NY, 1982.

[Ree92] Reed, Stephen K. *Cognition.* Third Edition. Brooks/Cole Publishing, Pacific Grove, CA, 1992.

[Rus95] Russell, Stuart and Norvig, Peter. *Artificial Intelligence A Modern Approach.* Prentice Hall,Upper Saddle River, NJ, 1995.

[Sau99] Sauter, Vicki L. "Intuitive Decision-Making" *Communications of the ACM.* Vol 24 No 6. ACM Publishing, New York, NY, June 1999.

[Sch96] Schultz, James V. *A Framework for Military Decision Making.* Masters Thesis, Air University, School of Advanced Airpower Studies, Maxwell AFB, AL, June 1996.

[Smi89] Smithson, Michael. *Ignorance and Uncertainty Emerging Paradigms.* Springer-Verlag, New York, NY, 1989.

[SML97] Schroeder, Will; Martin, Ken; and Lorensen, Bill. *The Visualization Toolkit,* Second Edition. Prentice Hall, Upper Saddle River, NJ, 1997.

[Tuf83] Tufte, Edward R. *Visual Display of Quantitative Information.* Graphics Press, Cheshire, CT, 1983.

[Tuf97] -----. *Visual Explanations.* Graphics Press, Cheshire, CT, 1997.

[TZ94] Tran, Tung and Zomorodi, Ali. "A Touch of Gray" *Vertices.* Duke University Undergraduate Publications, Durham, NC, 1994.

[Web97] Webster's Universal College Dictionary. Gramercy Books, New York, NY, 1997.

[WPL96] Wittenbrink, Craig M, Pang Alex T, and Lodha, Suresh K. "Glyphs for Visualizing Uncertainty in Vector Fields" *IEEE Transactions on Visualization and Computer Graphics.* IEEE Computer Society Press, San Jose, CA Vol 2 No 3, September 1996.

[WSSP99] DISA. *Common Warfighting Symbology MIL-STD 2525B.* DISA, Warrior Symbology Standardization Program, January 1999.

[Zim98] Zimmerman, Hans J. "Fuzzy Decision Support Systems" Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications. Springer, New York, NY, 1998.

Vita

Captain Evan T Watkins was born in Christiansted, Saint Croix in the US Virgin Islands. He graduated from St Croix Country Day School in June 1985. Evan left St Croix for the US Marine Corps in August and completed boot camp at Paris Island, South Carolina in November 1985. Following Infantry Training School he was stationed at the Naval Weapons Station Seal Beach, California for a little more than two years where he also met and married Stephany. His last year and a half were spent in the Fleet Marine Force with the 5th and 9th Marines at Camp Pendleton, California.

Evan returned to school after being Honorably Discharged from the Marine Corps in 1989 and earned a Bachelors of Science in Computer Science from the California State University, Dominguez Hills in December 1994. Returning to the military, Evan graduated from the US Air Force Officer Training School on 23 June 1995.

His first assignment was with the Operations Flight of the 61st Communications Squadron of the Space and Missile Center at Los Angeles AFB, California. During the three years at LA AFB Evan served as Information Systems manager, Network manager, Project manager, Chief of NCC daily operations, ... and finally as the Communication Squadron's Executive Officer.

Captain Watkins' follow-on assignment is with the Air Force Communications Agency, Air Force Communications and Information Center at Scott AFB, Illinois.

Permanent address: 18 Prospect Hill
Frederiksted, St. Croix
US Virgin Islands 00840

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
29 February 2000

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
Improving the Analyst and Decision-Maker's Perspective Through Uncertainty Visualization

5. FUNDING NUMBERS

6. AUTHOR(S)
Evan T. Watkins, Captain, USAF

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 P Street, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GCS/ENG/00M-24

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
NAIC/TAIT
Attn: Capt Allen Bobo
4180 Watson Way
Wright-Patterson AFB, OH 45433 DSN: 787-6327

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Dr Gregg H. Gunsch, ENG DSN: 785-6565, Ext. 4182

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

This thesis constructs the Taxonomy of Uncertainty and an approach for enhancing the information in decision support systems. The hierarchical categorization of numerous causes for uncertainty defines the taxonomy, which fostered the development of a technique for visualizing uncertainty. This technique is fundamental to expressing the multi-dimensional uncertainty that can be associated with any object. By including and intuitively expressing uncertainty, the approach facilitates and enhances intuition and decision-making without undue information overload.

The resulting approach for enhancing the information involves recording uncertainty, identifying the relevant items, computing and visualizing uncertainty, and providing interaction with the selection of uncertainty. A prototype embodying this approach to enhancing information by including uncertainty was used to validate these efforts. Evaluation responses of a small sample space support the thesis that the decision-maker's knowledge is enhanced with enlightening information afforded by including and visualizing uncertainty, which can improve the decision-making process.

Although the concept was initially conceived to help decision support system users deal with uncertainty, this methodology and these ideas can be applied to any problem where objects with many potential reasons for uncertainty are the focus of the decision-making.

14. SUBJECT TERMS
Uncertainty, Uncertainty Visualization, Gnomon, Uncertainty Taxonomy, Information Visualization, Decision-Making, Decision Analysis, Decision Support System, Fuzzy Logic

15. NUMBER OF PAGES
204

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL